

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw mode!

Run on:

December 30, 2003, 11:55:48 ; Search time 41 Seconds
(without alignments)

58.071 Million cell updates/sec

Title:

US-09-807-610-11

Score:

80

Sequence:

1 ALADLYBSEGCGGSGR 15

Scoring table:

BLOSUM62

Gapext 0.5

Gapop 10.0

Peptide score:

20

Sequence:

1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:

1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03;*

1: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1980 DAT:*

2: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA2001.DAT:*

23: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA2002.DAT:*

24: /SIDS1/gcadata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Query	Match	Length	DB ID	Description
1	80	100.0	15	21	AAY98151	Human icIL-IraII a
2	80	100.0	24	19	AAW65063	ICB binding peptid
3	80	100.0	180	17	AAR91361	Intracellular II-1
4	80	100.0	180	20	AAT32278	ICIRAP1 conserved
5	80	100.0	180	23	ABP52019	NOVINTRA A homolog
6	80	100.0	180	23	ABP2033	NOVINTRA B homolog
7	80	100.0	180	24	AB33566	Intracellular inter
8	72	90.0	19	19	AAW65061	ICB binding peptid
9	72	90.0	24	19	AAW65064	ICB binding peptid

A new expression vector for intracellular IL-1 receptor antagonist type

RESULT 1

AAV98151

standard; Protein: 15 AA.

AAV98151;

XX

DT 22-AUG-2000 (first entry)

XX

Human icIL-IraII amino acid fragment

XX

DE Human icIL-IraII

XX

Expression vector; icIL-IraII; interleukin-1; IL-1; growth hormone;

XX

signal peptide; intracellular IL-1 receptor antagonist type II; cytokine;

XX

human; IL-1 overexpression.

XX

OS Homo sapiens.

XX

PN WO200002146-A1.

PD 20-APR-2000.

XX

PF 14-OCT-1999;

XX

PR 14-OCT-1998;

XX

PA (INTE) INTERPHARM LAB LTD.

PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

PI Amitai H, Chitlaru E;

XX

DR WPI: 2000-317997/27.

XX

PT

PT II joined to a signal peptide is useful for reducing IL-1 in a patient
 PT and allows secretion of the intracellular molecule -
 PS Example 11; Page 17; 35pp; English.
 XX This invention relates to an expression vector comprising DNA encoding a
 CC signal peptide of a protein normally expressed and secreted by human
 CC cells (e.g. human growth hormone signal peptide) joined to DNA encoding
 CC intracellular interleukin 1 (IL-1) receptor antagonist type II
 CC (icIL-1ra), and operably linked to a promoter sequence. The expression
 CC results in a signal peptide fused in frame to icIL-1raII. Il-1ra is a
 CC unique IL-1 receptor antagonist, that inhibits signalling from the
 CC functional IL-1 receptor. icIL-1raII inhibits exogenous IL-1 dependent
 CC responses, and is termed intracellular due to its lack of signal
 CC peptide. The expression vector can be used in a process to engineer
 CC proteins to be secreted. The expression vector and resulting recombinant
 CC icIL-1raII protein are used to reduce IL-1 in a patient, either by direct
 CC administration of the glycosylated icIL-1raII, or by gene therapy using
 CC the vector. The present sequence analysis to prove that the expression vector
 CC obtained in protein sequence analysis to prove that the expression vector
 CC can be used to create secreted mature icIL-1ra type II protein.
 XX Sequence -15 AA;

Query Match 100.0%; Score 80; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db AAW65063 standard; Peptide; 24 AA.

Y 1 ALADLYEEGGGGGF 15

Db 1 ALADLYEEGGGGGF 15

:
 RESULT 2

AAW65063 standard; Peptide; 24 AA.

XX ARW65063;

XX 15-SBP-1998 (first entry)

XX ICB binding peptide C.

XX Interleukin-1-beta converting enzyme; ICB; autoimmune disease; infection;
 XX inflammatory diseases; rheumatoid arthritis; septic shock; leukaemia;
 XX multiple sclerosis; diagnosis; cysteine protease; inhibitor; cytokine.
 XX Synthetic.

XX Homo sapiens.

XX WO9818823-A1.

XX (ISTP) ARS APPLIED RES SYSTEMS HOLDING NV.

XX 07-MAY-1998.

XX 31-OCT-1996; 96WO-EP04738.

XX 31-OCT-1996; 96WO-EP04738.

XX (ISTP) ARS APPLIED RES SYSTEMS HOLDING NV.

XX Introna M, Mantovani A, Muzio M;

XX WPI; 1998-272138/24.

XX Peptide binding and inhibiting interleukin-1^b-converting enzyme.

PT ICB - useful in treating pathologies requiring inhibition of ICB or
 PT ICB family enzymes, e.g. autoimmune or inflammatory diseases etc.
 XX Claim 4; Figure 3; 32pp; English.
 XX AAW65060-W65067 are peptides capable of binding to interleukin-1-beta
 CC converting enzyme (ICB) and/or enzymes of the ICB family. Such peptides
 CC can be used medically, especially to prepare pharmaceutical compositions

CC which can be administered prophylactically or therapeutically to
 CC patients with pathologies requiring ICG inhibition and/or inhibition of
 CC enzymes of the ICB family, e.g. autoimmune diseases, lethal bacterial or
 CC viral infections or inflammatory diseases. Specific examples of such
 CC pathologies include rheumatoid arthritis, septic shock, leukaemia,
 CC multiple sclerosis etc. The peptide may also be used in the diagnosis of
 CC such pathologies. ICG is a heterodimeric cysteine protease which cleaves
 CC interleukin-1 beta precursor protein (pIL-1-beta) during interleukin-1
 CC beta (IL-1-beta) synthesis. Inhibition of ICG can thus inhibit synthesis
 CC of IL-1-beta. Interleukin-1's are potent inflammatory and pyrogenic
 CC cytokines that normally have beneficial effects but also act in many
 CC human pathologies, especially on the immune response and on inflammatory
 CC processes.

XX Sequence 24 AA;

Query Match 100.0%; Score 80; DB 19; Length 24;

Best Local Similarity 100.0%; Pred. No. 2.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy

1 ALADLYEEGGGGGF 15

2 ALADLYEEGGGGGF 16

RESULT 3

ARR91361

ID ARR91361 standard; Protein; 180 AA.

XX AC ARR91361;

XX DN 25-MAR-2003 (updated)

DT 22-Oct-1996 (first entry)

DE Intracellular IL-1 receptor antagonist type II.
 XX Intracellular IL-1 receptor antagonist; icIL-1ra;
 KW secreted IL-1 receptor antagonist; sIL-1ra;
 KW interleukin; IL-1a; IL-1B; auto-immune disease.
 XX Homo sapiens.

XX WO9612022-A1.

XX PN

PD 25-APR-1996.

PP 12-OCT-1995; 95WO-EP04023.

XX PR 13-OCT-1994; 94IT-MI02097.

XX PA (ISTP) ARS APPLIED RES SYST HOLDING NV.

XX PT Colotta F, Mantovani A, Muzio M;

XX WPI; 1996-222008/22.

DR DR N-PSDB; ARR15099.

XX PT IL-1 receptor antagonist active against IL-1a and -1B - for
 PT treating, preventing or diagnosing auto-immune diseases
 XX PS Claim 2; Page 22-23; 36pp; English.
 XX An new IL-1 receptor antagonist includes the sequence given in ARR91360.
 CC The complete icIL-1ra is given in ARR15099. The protein is
 CC expressed by DNA similar to that encoding the known receptor
 CC inhibitor icIL-1ra, but having a 63 bp insert between the first
 CC exon of icIL-1ra specific exon and the internal acceptor site of the first
 CC exon of sIL-1ra.
 CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 180 AA;

Query Match 100.0%; Score 80; DB 17; Length 180;

Query	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps	Score	Length	Method of invention
Qy	1 ALADLYEGGGGGGB 15	0.00023;	0;	0;	0;	100.0%	20;	PT
Db	2 ALADLYEGGGGGGB 16							PT
RESULT 4								
AY33278 standard; protein; 180 AA.								
ID	AY33278							
XR								
AC	AY33278;							
XX								
DR	23-NOV-1999 (first entry)							
XX								
DB	ICIRAPII conserved peptide fragment.							
XX								
KW	Interleukin-1; IL-1; cytokine; inhibitor; antagonist; receptor; trauma; anti-inflammatory; anti-infective; apoptotic; infection; catabolic; degeneration; chronic inflammation; autoimmune disease; IL-1b; nervous system; motor system; inhibitor; erythropoiesis; chondrocyte; rheumatic patient; collagen synthesis; hepatocyte; cell-cell adhesion; cartilage degradation; nerve regeneration; necrosis induction; apoptosis; intracellular IL-1 receptor agonist; ICIRAPII.							
XX								
OS	Unidentified.							
XX								
PN	DB19753753-A1.							
XX								
PD	26-AUG-1999.							
XX								
PP	04-DEC-1997; 97DE-1053753.							
XX								
PR	04-DEC-1997; 97DE-1053753.							
XX								
PA	(ÖRTH-) ORTHGEN GENTECHNOLOGIE GMBH.							
XX								
PI	Meijer H, Wehling P, Reinecke J;							
XX								
DR	WPI; 1999-479917/41.							
XX								
PT	New recombinant nucleic acid encoding proteins for modifying effects of cytokine systems, e.g. for treating infection or trauma - Disclosure, Fig 4; 8pp; German.							
XX								
PS								
XX								
CC	This invention describes novel recombinant nucleic acids (I) that encode proteins (II) which modify the effects of cytokines or their inhibitors/antagonists or their receptors. The products of the invention (II) have anti-inflammatory, anti-infective and apoptotic activity. (I) and (II) are particularly used to modulate activities of the human interleukin-1 (IL-1) system which is involved in many inflammatory and catabolic processes, e.g. in infection, trauma, degeneration and chronic inflammation (autoimmune diseases) of the nervous and motor systems and internal organs. In particular, (II) are used in cases where IL-1 inhibits erythropoiesis in rheumatic patients, inhibits collagen synthesis by chondrocytes, inhibits growth of hepatocytes, promotes cell-cell adhesion, immune reactions, degradation of cartilage, nerve regeneration, and induction of necrosis or apoptosis in the central nervous system. This sequence represents a fragment of an intracellular IL-1 receptor agonist (ICIRAPII) which is used to describe the method of the invention.							
CC	Sequence 180 AA;							
CC	Query Match Score 80; Pred. No. 0.00023; Mismatches 0; Indels 0; Gaps 0;							
CC	Best Local Similarity 100.0%; Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	1 ALADLYEGGGGGB 15							
Db	2 ALADLYEGGGGGB 16							

CC IL-1 epsilon protein is useful for treating a lung disease such as lung
 CC cancer, asthma, emphysema, allergic lung irritation and lung inflammation
 CC in a mammal. ABQ7396 to ABQ7027 and ABP51961 to ABP5048 represent
 CC sequences used in the exemplification of the present invention.

Sequence	180 AA;	Query Match	100.0%;	Score	80;	DB	23;	Length	180;
		Best Local Similarity	100.0%;	Pred. No.	0.00023;				
		Matches	15;	Mismatches	0;	Indels	0;	Gaps	0;
Qy		1	ALADLYBEGGGGGGB	15					
Dy		2	ALADLYBEGGGGGGE	16					

RESULT 6
 ABP5033 ABP52033 standard; Protein; 180 AA.
 XX AC ABP52033;
 XX DT 10-OCT-2002 (first entry)
 DB NOVINTRA B homologous amino acid sequence SEQ ID NO:79.
 XX KW transmembrane protein; neuromedin protein; gonadotropin protein;
 KW interleukin-1 receptor antagonist; interleukin-1 epsilon; NOVX;
 KW IL-1 epsilon; IL-1 receptor antagonist; lung disease; nootropic;
 KW cytostatic; neuroprotective; antiinflammatory; antibacterial;
 KW immunosuppressive; cerebroprotective; antidiabetic; antiarthritic;
 KW antiasthmatic; antiallergic; gene therapy; antibody-based therapy;
 KW cell signalling disorder; haematopoietic disorder; endocrine; muscle;
 KW neurodegenerative disorder; neurological disorder; cancer; melanoma;
 KW central nervous system cancer; reproductive development disorder;
 KW metabolic function disorder; bone metabolism; structure disorder; stroke;
 KW inflammatory response disorder; immune regulation disorder; septic shock;
 KW diabetes; arthritis; lung cancer; emphysema; allergic lung irritation;
 KW lung inflammation.
 OS Homo sapiens.
 XX PN US2002068279-A1.
 XX PD 06-JUN-2002.
 XX PP 05-DEC-2000; 2000US-0730617.
 XX PR 06-DEC-1999; 99US-169056P.
 XX PR 09-DEC-1999; 99US-169866P.
 XX PR 09-DEC-1999; 99US-169868P.
 XX PR 10-DEC-1999; 99US-170252P.
 XX PR 12-JAN-2000; 2000US-175740P.
 XX PA (CURA-) CORAGEN CORP.
 XX PI Burgess C, Prayaga SK, Shimkets RA, Rastelli L, Zerhusen B;
 PI Mezes P;
 XX DR WPI; 2002-58242/62.

CC antiallergic activities, and can be used in gene therapy and antibody-
 CC based therapy. NOVX polypeptides, nucleic acid (I) encoding them and an
 CC antibody (III) that binds the polypeptide, are useful for treating or
 CC preventing a NOVX protein-associated disorder in humans. NOVTRAN can be
 CC used in the treatment of a cell signalling disorder, such as, a
 CC haematopoietic disorder or a neurodegenerative disorder. NOVTR can be
 CC used in the treatment of an endocrine, muscle, neurological disorder,
 CC central nervous system cancer, breast, colon, ovarian, kidney, prostate
 CC or thyroid cancer. NOVCON can be used in the treatment of a reproductive
 CC development disorder, metabolic function disorder or melanoma. NOVNTRA
 CC proteins can be used in the treatment of and a bone metabolic or
 CC structure disorder, an inflammatory response disorder, an immune
 CC regulation disorder, septic shock, stroke, diabetes, arthritis or
 CC cancer. An agent which modulates the expression or activity of a human
 CC IL-1 epsilon protein is useful for treating a lung disease such as lung
 CC cancer, asthma, emphysema, allergic lung irritation and lung inflammation
 CC in a mammal. ABQ7396 to ABQ74027 and ABP51981 to ABP52048 represent
 CC sequences used in the exemplification of the present invention.

Query Match	100.0%;	Score	80;	DB	23;	Length	180;		
		Best Local Similarity	100.0%;	Pred. No.	0.00023;				
		Matches	15;	Mismatches	0;	Indels	0;	Gaps	0;
Qy		1	ALADLYBEGGGGGB	15					
Dy		2	ALADLYBEGGGGGC	16					

RESULT 7
 ABAB3566 standard; Protein; 180 AA.
 ID AAB33566
 AC AAB33566;
 XX DT 16-APR-2003 (first entry)
 DB Intracellular interleukin-1RN (IL-1RN).
 XX KW drug screening; fungicide; gene therapy; antibacterial; infection;
 KW virucide; interleukin-1RN; IL-1RN.
 OS Unidentified.
 XX PN WO2002101015-A2.
 XX PR 19-DBC-2002.
 XX PR 11-JUN-2002; 2002WO-US183446.
 XX PR 11-JUN-2001; 2001US-297305P.
 XX PA (INTB-) INTERLEUKIN GENETICS INC.
 XX PI Dower S, Duff GW;
 XX DR WPI; 2003-148793/14.
 XX PR New detection reagent, useful for monitoring molecular assembly events
 PR to permit the dissection of genetic and non-genetic influences on
 PR biological activity, comprises an interactive sensor pair -
 XX PS Claim 17; Fig 8; 56pp; English.
 XX PR The invention relates to methods, compositions and apparatus for
 CC monitoring molecular assembly events. It also relates to a detection
 CC reagent comprising an interactive sensor pair. The detection reagent is
 CC useful for monitoring molecular assembly events to permit the dissection
 CC of genetic and non-genetic influences on a particular biological
 CC activity. The method is useful for linking genetic variations to
 CC molecular and physiological events, drug screening, diagnostics, therapy
 CC selection and dosing, patient monitoring or environmental safety. The

Db 2 ALAALYEEGGGGGG 16

RESULT 10
 AAR91360
 ID AAR91360 standard; Peptide; 21 AA.
 XX AC AAR91360;
 XX DT 25-MAR-2003 (updated)
 DT 22-OCT-1996 (first entry)
 XX DE Intracellular IL-1 receptor antagonist type II fragment.
 XX KW Intracellular IL-1 receptor antagonist; icIL-1ra;
 KW secreted IL-1 receptor antagonist; sIL-1ra;
 KW interleukin; IL-1a; IL-1B; auto-immune disease.
 XX DS Homo sapiens.
 PN WO9512022-A1.
 XX PD 25-APR-1996.
 XX PP 12-OCT-1995; 95WO-BP04023.
 XX PR 13-OCT-1994; 94IT-MI02097.
 XX PA (ISPF) ARS APPLIED RES SYSTEM HOLDING NV.
 PI Colotta F, Mantovani A, Muzio M;
 XX DR WPI; 1995-222008/22.

PT IL-1 receptor antagonist active against IL-1a and -1B - for
 PT treating, preventing or diagnosing auto-immune diseases
 PS Claim 1; Page 21; 36pp; English.

XX An new IL-1 receptor antagonist includes the sequence given in AAR91360.
 XX The complete icIL-1ra is given in AAR91360. The protein is
 XX expressed by DNA similar to that encoding the known receptor
 XX inhibitor icIL-1ra, but having a 63 bp insert between the first
 XX exon of icIL-1ra specific exon and the internal acceptor site of the first
 XX exon of sIL-1ra.
 XX (Updated on 25-MAR-2003 to correct PR field.)

SQ Sequence 21 AA;

Query Match 83.8%; Score 67; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 3 ADLYEEGGGGG 14
 | ||| | | | | | | | |
 1 ADLYEEGGGGG 12

RESULT 11
 AAW65060
 ID AAW65060 standard; peptide; 19 AA.
 XX AC AAW65060;
 XX DT 15-SBP-1998 (first entry)
 XX IC binding peptide #1.

XX Interleukin-1-beta converting enzyme; ICB; autoimmune disease; infection;
 XX inflammatory disease; rheumatoid arthritis; septic shock; leukaemia;
 XX multiple sclerosis; diagnosis; cysteine protease; inhibitor; cytokine.
 XX Synthetic.
 OS Homo sapiens
 XX

OS Homo sapiens

XX Key Location/Qualifiers

XX Misc-difference 2

FT /note= "can be either Ala or more preferably Asp"
 FT Misc-difference 19
 FT /note= "can be either Ala or more preferably Asp"

XX PN WO9818923-A1.

XX PD 07-MAY-1998.

XX PP 31-OCT-1996; 96WO-BP04738.

XX PR 31-OCT-1996; 96WO-BP04738.

XX PA (ISPF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PI Introna M, Mantovani A, Muzio M;
 XX DR WPI; 1998-272138/24.

XX Peptide binding and inhibiting interleukin-1 β -converting enzyme,
 PT ICB - useful in treating pathologies requiring inhibition of ICB or
 PR ICB family enzymes, e.g. autoimmune or inflammatory diseases etc.

PS Claim 1; Page 14; 32pp; English.

XX AAW65060-W65067 are peptides capable of binding to interleukin-1-beta
 CC converting enzyme (ICB) and/or enzymes of the ICB family. Such peptides
 CC can be used medicinally, especially to prepare pharmaceutical compositions
 CC which can be administered prophylactically or therapeutically to
 CC patients with pathologies requiring ICB inhibition and/or inhibition of
 CC enzymes of the ICB family, e.g. autoimmune diseases, bacterial or
 CC viral infections or inflammatory diseases. Specific examples of such
 CC pathologies include rheumatoid arthritis, septic shock, leukemia,
 CC multiple sclerosis etc. The peptide may also be used in the diagnosis of
 CC such pathologies. ICB is a heterodimeric cysteine protease which cleaves
 CC interleukin-1 beta precursor protein (pIL-1-beta) during interleukin-1
 CC beta (IL-1-beta) synthesis. Inhibition of ICB can thus inhibit synthesis
 CC of IL-1-beta. Interleukin-1's are potent inflammatory and pyrogenic
 CC cytokines that normally have beneficial effects but also act in many
 CC human pathologies, especially on the immune response and on inflammatory
 CC processes.

XX Sequence 19 AA;

Query Match 81.2%; Score 65; DB 19; Length 19;
 Best Local Similarity 92.3%; Pred. No. 0.0036;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ADLYEEGGGG 15
 | ||| | | | | | | | |
 1 AXLYEEGGGG 13

DB

RESULT 12

AAW65062 AAW65062 standard; peptide; 19 AA.

XX AAW65062;

XX DR 15-SBP-1998 (first entry)

XX ICB binding peptide #3.

XX Interleukin-1-beta converting enzyme; ICB; autoimmune disease; infection;
 XX inflammatory disease; rheumatoid arthritis; septic shock; leukaemia;
 XX multiple sclerosis; diagnosis; cysteine protease; inhibitor; cytokine.
 XX Synthetic.
 OS Homo sapiens
 XX

PN	WO9418345-A1.
XX	
PD	07-MAY-1998.
XX	
PP	31-OCT-1996; 96WO-EP04738.
XX	
PR	31-OCT-1996; 96WO-EP04738.
XX	
PA	(ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
XX	
PI	Intronca M, Mantovani A, Muzio M;
XX	
DR	WPI; 1998-272138/24.
XX	
PT	Peptide binding and inhibiting interleukin-1 ^b -converting enzyme.
PT	ICB - useful in treating pathologies requiring inhibition of ICB or
PT	ICB family enzymes, e.g. autoimmune or inflammatory diseases etc.
XX	
PS	Claim 3; Page 14; 32pp; English.
XX	
CC	AW65060-W65067 are peptides capable of binding to interleukin-1-beta converting enzyme (ICB) and/or enzymes of the ICB family. Such peptides can be used medicinally, especially to prepare pharmaceutical compositions which can be administered prophylactically or therapeutically to patients with pathologies requiring ICB inhibition and/or inhibition of enzymes of the ICB family, e.g. autoimmune diseases, lethal bacterial or viral infections or inflammatory diseases. Specific examples of such pathologies include rheumatoid arthritis, septic shock, leukaemia, multiple sclerosis etc. The peptide may also be used in the diagnosis of such pathologies. ICB is a heterodimeric cysteine protease which cleaves interleukin-1 beta precursor protein (pIL-1-beta) during interleukin-1 beta (IL-1-beta) synthesis. Inhibition of ICB can thus inhibit synthesis of IL-1-beta. Interleukin-1's are potent inflammatory and pyrogenic cytokines that normally have beneficial effects but also act in many human pathologies, especially on the immune response and on inflammatory processes.
CC	
CC	AW65060-W65067 are peptides capable of binding to interleukin-1-beta converting enzyme (ICB) and/or enzymes of the ICB family. Such peptides can be used medicinally, especially to prepare pharmaceutical compositions which can be administered prophylactically or therapeutically to patients with pathologies requiring ICB inhibition and/or inhibition of enzymes of the ICB family, e.g. autoimmune diseases, lethal bacterial or viral infections or inflammatory diseases. Specific examples of such pathologies include rheumatoid arthritis, septic shock, leukaemia, multiple sclerosis etc. The peptide may also be used in the diagnosis of such pathologies. ICB is a heterodimeric cysteine protease which cleaves interleukin-1 beta precursor protein (pIL-1-beta) during interleukin-1 beta (IL-1-beta) synthesis. Inhibition of ICB can thus inhibit synthesis of IL-1-beta. Interleukin-1's are potent inflammatory and pyrogenic cytokines that normally have beneficial effects but also act in many human pathologies, especially on the immune response and on inflammatory processes.
XX	
SQ	Sequence 19 AA;
XX	
Query Match	Score 64; DB 19; Length 19;
Best Local Similarity	90.0%; Pred. No. 0.0051;
Matches 12;	Gaps 0;
Conservative 0;	Mismatches 1;
Indels 0;	
Qy	3 ADLYBGGGGGG 15
Db.	1 ADLYBGGGGGG 13
XX	
RESULT 13	
ID	AAR60407
ID	AAR60407 standard; peptide; 24 AA.
XX	
AC	AAR60407;
XX	
DT	25-MAR-2003 (updated)
DT	30-MAR-1995 (first entry)
XX	
DB	Antiproliferative peptide C (repeat) to human B-cell lymphoma.
XX	
KW	antiproliferative; transplant; B-cell lymphoma line SUP-B8; Burkitt's;
KW	inhibit clonal expansion; induce apoptosis; anti-idiotype; IgM lambda;
KW	inhibit cell proliferation; peptidomimetics; cell surface receptor;
KW	immunoglobulin superfamily; transcription; neoplasia; identification;
KW	induce replication; therapy; clonal anergy; modulate tyrosine kinase.
XX	
OS	synthetic.
XX	
Key	Location/Qualifiers
PH	Misc-difference 1..9 /note= "direct repeat"
FT	Misc-difference 1..9 /note= "direct repeat"
FT	FT 16..24 /note= "direct repeat"
XX	
PN	W09418345-A1.
XX	
PD	18-AUG-1994.
XX	
PR	04-FEB-1994;
XX	
PR	93US-0014426.
XX	
PR	15-NOV-1993;
XX	
PA	(AFFY-) AFFYMAX TECHNOLOGIES NV.
PA	(STRD) UNIV LELAND STANFORD JUNIOR.
XX	
PI	Bhatt RR, Dower WJ, Levy R, Renachler MF;
XX	
DR	WPI; 1994-279762/34.
XX	
PT	Identifying anti-proliferative peptide(s) which specifically bind to immunoglobulin super-family species idiotype - esp. to inhibit B-cell lymphoma and leukocytic leukaemia cell proliferation, for anti-idiotype therapy
PT	
PT	Claim 7; Page 49; 69pp; English.
XX	
CC	AARE0400-73 are peptide ligands which bind to purified IgM lambda receptor of the human Burkitt's lymphoma cell line SUP-B8. The peptides were identified with the use of filamentous phage libraries displaying random peptides. Corresponding synthetic peptides bound specifically to this Ig receptor, and blocked the binding of an anti-idiotype antibody. The ligands, when conjugated to form dimers or tetramers, induced cell death by apoptosis in vitro at nanomolar concentrations. This effect was associated with the specific stimulation of intracellular protein tyrosine phosphorylation. The peptides of the invention can be used individually, as complexes of cross-linked peptides or can be conjugated to deliver toxins or radionuclides to neoplastic cells bearing the specific Ig receptor.
CC	
CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	
SQ	Sequence 24 AA;
XX	
Query Match	Score 53; DB 15; Length 24;
Best Local Similarity	81.8%; Pred. No. 0.29;
Matches 9;	Mismatches 0;
Conservative 9;	Indels 2;
Qy	4 DLYBEGGGGG 14
Db	5 DLTRGGGGGG 15
XX	
RESULT 14	
ID	ABB29117
ID	ABB29117 standard; Peptide; 48 AA.
XX	
AC	ABB29117;
XX	
XX	Human; microarray; single exon probe; gene expression; breast; disease; cancer.
XX	
DT	01-FEB-2002 (first entry)
XX	
DS	Peptide #1768 encoded by breast cell single exon nucleic acid probe.
XX	
XX	W0200157271-A2.
XX	
PD	09-AUG-2001.
XX	
PP	30-JAN-2001; 2001WO-US006652.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-NOV-2000; 2000US-0608408.
PR	30-PER-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024233.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR; WPI; 2001-496933/54.
 XX PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 PS Claim 27; SEQ ID NO 12085; 327pp + sequence listing; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosis, breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 SQ Sequence 48 AA;

Query Match 63.7%; Score 51; DB 22; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 EBGGGGGB 15
 Db 6 EBGGGGGB 14

Search completed: December 30, 2003, 12:01:34
 Job time : 42 secs

RESULT 15
 AAM55071
 ID AAM55071 standard; Protein; 48 AA.
 XX DT 05-NOV-2001 (first entry)
 XX DB Human brain expressed single exon probe encoded protein SEQ ID NO: 27176.
 KW Human; brain expressed exon; gene expression analysis; Probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX OS Homo sapiens.
 PN WO200157275-A2.
 XX PD 09-AUG-2001.
 XX PP 30-JAN-2001; 2001WO-US00667.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.

Db • 2 ALADLYBEGGGGGE 16

RESULT 2
US-09-730-617-79
Sequence 79, Application US/09730617

Patent No. US20020068229A1

GENERAL INFORMATION:

APPLICANT: Burgess, Catherine B

APPLICANT: Pravaga, Sudhiradas K

APPLICANT: Shimbets, Richard A

APPLICANT: Rastelli, Luca

APPLICANT: Zerhusen, Bryan D

APPLICANT: Mezes, Peter S

TITLE OF INVENTION: *Nuc. US20020068279A1* Proteins and Nucleic Acids Encoding the Same

FILE REFERENCE: 15966-609

CURRENT APPLICATION NUMBER: US/09/730, 617

CURRENT FILING DATE: 2000-12-05

PRIOR APPLICATION NUMBER: 60/169, 056

PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: 60/169, 886

PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 60/169, 866

PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 60/170, 252

PRIOR FILING DATE: 1999-12-10

PRIOR APPLICATION NUMBER: 60/175, 740

PRIOR FILING DATE: 2000-01-12

NUMBER OF SEQ ID NOS: 100

SEQ ID NO: 79

LENGTH: 180

ORGANISM: Homo sapiens

US-09-730-617-79

Query-Match 100 0%; Score 80; DB 9; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALADLYBEGGGGGE 15
Db 2 ALADLYBEGGGGGE 16

RESULT 3
US-10-167-127-8
Sequence 8, Application US/10167127

Publication No. US20030100031A1

GENERAL INFORMATION:

APPLICANT: DOWER, STEVEN

APPLICANT: DUFF, GORDON W.

TITLE OF INVENTION: INTEGRATIVE ASSAYS FOR MONITORING MOLECULAR ASSEMBLY

TITLE OF INVENTION: EVENTS

FILE REFERENCE: MSA-026-01 (20974-2601)

CURRENT APPLICATION NUMBER: US/10/167,127

CURRENT FILING DATE: 2002-06-11

PRIOR APPLICATION NUMBER: 60/297,305

PRIOR FILING DATE: 2001-06-11

NUMBER OF SEQ ID NOS: 32

SEQ ID NO: 8

LENGTH: 180

ORGANISM: Homo sapiens

US-10-167-127-8

Query-Match 100 0%; Score 80; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALADLYBEGGGGGE 15

Db ||||||| 2 ALADLYBEGGGGGE 16

RESULT 4
US-10-029-386-32712
Sequence 32712, Application US/10029386

Publication No. US2003019704A1

GENERAL INFORMATION:

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Penn, Sharron G.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO

FILE REFERENCE: AECOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029, 386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO: 32712

LENGTH: 29

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC007759.1

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7

US-10-029-386-32712

Query Match 63.7%; Score 51; DB 12; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.2%;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 BEGGGGGG 15

Db 20 BEGGGGGG 28

RESULT 5
US-09-864-761-42759
Sequence 42759, Application US/09864761

Publication No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Penn, Sharron G.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: AECOMICA-X-1

CURRENT APPLICATION NUMBER: US/09/864, 761

CURRENT FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/180, 312

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 60/207, 456

PRIOR FILING DATE: 2000-09-03

PRIOR APPLICATION NUMBER: GB 24263, 6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/235, 359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00656

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00657

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00654

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00659

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR APPLICATION NUMBER: PCT/US01/00671

PRIOR APPLICATION NUMBER: PCT/US01/00672

PRIOR APPLICATION NUMBER: PCT/US01/00673

Db ||||||| 2 ALADLYBEGGGGGE 16

Qy 1 ALADLYBEGGGGGE 15

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006688
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006683
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006660
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006659
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006658
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006653
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006652
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006651
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006650
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 42759
 LENGTH: 48
 TYPE: PRT
 FEATURE:
 OTHER INFORMATION: MAP TO AC009664.3
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
 NAME/KEY: unsure
 LOCATION: 45
 US-09-864-761-42759
 Query Match 63.7%; Score 51; DB 9; Length 48;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 Qy 7 BGGGGGGG 15
 Db 6 BGGGGGGG 14
RESULT 6
 US-09-864-761-35634
 / Sequence 35634, Application US/09864761
 / Patent No. US20020048763A1
 / GENERAL INFORMATION:
 / APPLICANT: Penn, Sharron G.
 / APPLICANT: Rank, David R.
 / APPLICANT: Hanzel, David K.
 / APPLICANT: Chen, Wenheng
 / TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 / TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 / PILE UP REFERENCE: Aeomica-X-1
 / CURRENT APPLICATION NUMBER: US/09/864,761
 / CURRENT FILING DATE: 2001-05-23
 / PRIOR APPLICATION NUMBER: US 60/180,312
 / PRIOR FILING DATE: 2000-02-04
 / PRIOR APPLICATION NUMBER: US 60/207,456
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: US 09/632,366
 / PRIOR FILING DATE: 2000-08-03
 / PRIOR APPLICATION NUMBER: GB 24263.6
 / PRIOR FILING DATE: 2000-10-04
 / PRIOR APPLICATION NUMBER: US 60/236,359
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: PCT/US01/00666
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/006667
 / PRIOR APPLICATION NUMBER: PCT/US01/006664
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/006669
RESULT 7
 US-10-322-892-3
 / Sequence 3, Application US/10322892
 / Publication No. US20030171257A1
 / GENERAL INFORMATION:
 / APPLICANT: STIRBLE, ROBERT C.
 / APPLICANT: STIRBLE, MALCOLM L.
 / APPLICANT: XU, JIMMY
 / APPLICANT: VITETTA, ELLEN S.
 / APPLICANT: WILK, PETER J.
 / TITLE OF INVENTION: METHOD AND RELATED COMPOSITION EMPLOYING NANOSTRUCTURES
 / FILE REFERENCE: WO7-505
 / CURRENT APPLICATION NUMBER: US/10/322,892
 / CURRENT FILING DATE: 2002-12-18
 / PRIOR APPLICATION NUMBER: 60/342,894
 / PRIOR FILING DATE: 2001-12-19
 / NUMBER OF SEQ ID NOS: 4
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 3
 / LENGTH: 258
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-322-892-3
 Query Match 63.7%; Score 51; DB 12; Length 258;

Best Local Similarity 81.8%; Pred. No. 18;
Matches 9; Conservative 1; Mismatches 1;
Indels 0; Gaps 0;

Qy 4 LYBEGGGGGG 14
Db 122 DNWEEGGGGG 132

RESULT 8

US-09-949-192-23
Sequence 23, Application US/09449192
Patent No. US20020142232A1
GENERAL INFORMATION:
APPLICANT: Parham, Christi L.
APPLICANT: Gorman, Daniel L.
APPLICANT: Kurata, Hirokazu
APPLICANT: Arai, Naoko
APPLICANT: Sana, Theodore R.
APPLICANT: Mattson, Jeanine D.
APPLICANT: Murphy, Brian B.
APPLICANT: Savkoor, Chetan
APPLICANT: Grain, Jeffrey
APPLICANT: Smith, Kathleen M.
APPLICANT: McClanahan, Terrill K.
TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS AND METHODS
FILE REFERENCE: DX01169K
CURRENT APPLICATION NUMBER: US/09/949,192
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/231,,267
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 23
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-949-192-23

Query Match 62.5%; Score 50; DB 10; Length 977;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 8; Conservative 2; Mismatches 0;
Indels 0; Gaps 0;

Qy 5 LYBEGGGGGG 14
Db 708 LPEDGGGGGG 717

RESULT 9

US-10-190-115-14
Sequence 34, Application US/10190115
Publication No. US20030207394A1
GENERAL INFORMATION:
APPLICANT: Alsobrook, John P. II
APPLICANT: Boldog, Ferenc L.
APPLICANT: Burgess, Catherine B.
APPLICANT: Casman, Stacie J.
APPLICANT: Gross, William M.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Ji, Weizhen
APPLICANT: Lepley, Denise M.
APPLICANT: Liu, Xiaohong
APPLICANT: Mezick, Amanda J.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Rastelli, Luca
APPLICANT: Shen, Lei
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shimbets, Richard A.
APPLICANT: Spydak, Kimberly A.
APPLICANT: Szekeres, Edward S. Jr.
APPLICANT: Taupier, Raymond J. Jr.
APPLICANT: Tchernev, Velizar T.

APPLICANT: Zerhusen, Bryan D.
APPLICANT: Voss, Edward Z.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLIC ACIDS ENCODING SAME
CURRENT APPLICATION NUMBER: US/10/190,115
CURRENT FILING DATE: 2003-02-10
PRIOR APPLICATION NUMBER: 60/1303,168
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/1386,816
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/1368,996
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/1215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/1215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/1215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/1216,585
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/1216,586
PRIOR FILING DATE: 2001-07-07
PRIOR APPLICATION NUMBER: 60/1216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/1218,622
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 136
SOFTWARE: CurateSeqList version 0.1
SEQ ID NO: 34
LENGTH: 977
TYPE: PRT
ORGANISM: Homo sapiens
US-10-190-115-34

Query Match 62.5%; Score 50; DB 12; Length 977;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LYBEGGGGGG 14
Db 708 LPEDGGGGGG 717

RESULT 10

US-10-424-233-21
Sequence 21, Application US/10424233
Publication No. US20030220263A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: NOVEL HUMAN LEUCINE-RICH REPEAT-CONTAINING PROTEINS SPECIFICALLY EXPRESSED IN THE NERVOUS SYSTEM
TITLE OF INVENTION: EXPRESSED IN THE NERVOUS SYSTEM
FILE REFERENCE: D0233.NP
CURRENT APPLICATION NUMBER: US/10/424,233
CURRENT FILING DATE: 2003-04-25
PRIOR APPLICATION NUMBER: U.S. 60/375,335
PRIOR FILING DATE: 2002-04-25
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 21
LENGTH: 977
TYPE: PRT
ORGANISM: Homo sapiens
US-10-424-233-21

Query Match 62.5%; Score 50; DB 12; Length 977;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LYBEGGGGGG 14
Db 708 LPEDGGGGGG 717

RESULT 11
 US-09-934-455-138
 / Sequence 138, Application US/09934455
 / Publication No. US20030121070A1
 / GENERAL INFORMATION:
 / APPLICANT: Creeiman, Robert
 / APPLICANT: Dubell, Arnold T.
 / APPLICANT: Heard, Jacqueline
 / APPLICANT: Jiang, Cai-Zhong
 / APPLICANT: Keddie, James
 / APPLICANT: Pilgrim, Marsha
 / APPLICANT: Ratcliffe, Oliver
 / APPLICANT: Reuber, Lynne
 / APPLICANT: Riechmann, Jose Luis
 / APPLICANT: Yu, Guo-Liang
 / APPLICANT: Pineda, Omaira
 / TITLE OF INVENTION: Genes for Modifying Plant Traits IV
 / FILE REFERENCE: MBI-0025
 / CURRENT APPLICATION NUMBER: US/09/934,455
 / CURRENT FILING DATE: 2001-08-22
 / PRIOR APPLICATION NUMBER: 60/227439
 / PRIOR FILING DATE: 2000-08-22
 / PRIOR APPLICATION NUMBER: MBI-0022
 / PRIOR FILING DATE: 2001-11-16
 / PRIOR APPLICATION NUMBER: MBI-0023
 / PRIOR FILING DATE: 2001-04-17
 / NUMBER OF SEQ ID NOS: 516
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 934-455-138
 / LENGTH: 325
 / TYPE: PRT
 / ORGANISM: Arabidopsis thaliana
 / SEQ ID NO 934-455-138

Query 2 LADLYBEGGGGG 15
 Db 264 LASGTEGGGGGB 277

Query Match 60.0%; Score 48; DB 11; Length 325;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 10; Conservative 0; Mismatches 4; Indels 0;
 Gaps 0;

RESULT 12
 US-10-225-068-246
 / Sequence 246, Application US/10225068
 / Publication No. US200301217383A1
 / GENERAL INFORMATION:
 / APPLICANT: Mendel Biotechnology, Inc.
 / APPLICANT: Reuber, T. Lynne
 / APPLICANT: Riechmann, Jose Luis
 / APPLICANT: Heard, Jacqueline E.
 / APPLICANT: Jiang, Cai-Zhong
 / APPLICANT: Adam, Luc J.
 / APPLICANT: Dubell, Arnold T.
 / APPLICANT: Ratcliffe, Oliver
 / APPLICANT: Pineda, Omaira
 / APPLICANT: Yu, Guo-Liang
 / APPLICANT: Broun, Pierre E.
 / TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
 / TITLES OF INVENTION: POLYPEPTIDES IN PLANTS
 / FILE REFERENCE: 51442002040
 / CURRENT APPLICATION NUMBER: US/10/225,068
 / CURRENT FILING DATE: 2000-08-09
 / PRIOR APPLICATION NUMBER: 60/310,847
 / PRIOR FILING DATE: 2001-08-09
 / PRIOR APPLICATION NUMBER: 60/336,049
 / PRIOR FILING DATE: 2001-11-19
 / PRIOR APPLICATION NUMBER: 60/338,692
 / PRIOR FILING DATE: 2001-12-11
 / PRIOR APPLICATION NUMBER: 10/171,468

Query Match 60.0%; Score 48; DB 12; Length 325;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 10; Conservative 0; Mismatches 4; Indels 0;
 Gaps 0;

Query 2 LADLYBEGGGGG 15
 Db 264 LASGTEGGGGGB 277

Query Match 60.0%; Score 48; DB 12; Length 325;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 10; Conservative 0; Mismatches 4; Indels 0;
 Gaps 0;

RESULT 13
 US-10-302-267-62
 / Sequence 62, Application US/10302267
 / Publication No. US20030229915A1
 / GENERAL INFORMATION:
 / APPLICANT: Keddie, James
 / APPLICANT: Fromm, Michael
 / APPLICANT: Heard, Jacqueline
 / APPLICANT: Riechmann, Jose Luis
 / APPLICANT: Adam, Luc
 / APPLICANT: Broun, Pierre
 / APPLICANT: Pineda, Omaira
 / APPLICANT: Reuber, Lynne
 / APPLICANT: Zhang, James
 / APPLICANT: Yu, Guo-Liang
 / APPLICANT: Jiang, Cai-Zhong
 / APPLICANT: Samaha, Raymond
 / APPLICANT: Pilgrim, Marsha
 / APPLICANT: Creeiman, Robert
 / TITLE OF INVENTION: PLANT GENE SEQUENCES II
 / FILE REFERENCE: MBI-0007
 / CURRENT APPLICATION NUMBER: US/10/302,267
 / PRIOR APPLICATION NUMBER: US/09/506,720
 / PRIOR FILING DATE: 2002-11-22
 / PRIOR APPLICATION NUMBER: 60/120,880
 / PRIOR FILING DATE: 1999-02-18
 / PRIOR APPLICATION NUMBER: 60/121,037
 / PRIOR FILING DATE: 1999-02-22
 / PRIOR APPLICATION NUMBER: 60/124,278
 / PRIOR FILING DATE: 1999-03-11
 / PRIOR APPLICATION NUMBER: 60/129,450
 / PRIOR FILING DATE: 1999-04-15
 / PRIOR APPLICATION NUMBER: 60/135,134
 / PRIOR FILING DATE: 1999-05-20
 / PRIOR APPLICATION NUMBER: 60/144,153
 / PRIOR FILING DATE: 1999-07-15
 / PRIOR APPLICATION NUMBER: 60/161,143
 / PRIOR FILING DATE: 1999-10-22
 / PRIOR APPLICATION NUMBER: 60/162,656
 / PRIOR FILING DATE: 1999-11-01
 / NUMBER OF SEQ ID NOS: 218
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 62
 / LENGTH: 325
 / TYPE: PRT
 / ORGANISM: Arabidopsis thaliana
 / FEATURE:
 / OTHER INFORMATION: G805
 / US-10-302-267-62

Query Match 60.0%; Score 48; DB 12; Length 325;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 .LADLYBEGGGGGGB 15
 Db 264 LASGYGGGGGGGB 277

RESULT 14
 US-09-815-242-10447
 Sequence 10447, Application US/09815242
 / Patent No. US2002006156A1
 / GENERAL INFORMATION:
 / APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Kari L.
 / APPLICANT: Zyskind, Judith W.
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John D.
 / APPLICANT: Carr, Grant J.
 / APPLICANT: Yamamoto, Robert T.
 / APPLICANT: Xu, H. Howard
 / TITLE OF INVENTION: Identification of Essential Genes in
 / Prokaryotes
 / FILE REFERENCE: ELITRA.011A
 / CURRENT APPLICATION NUMBER: US/09/815,242
 / CURRENT FILING DATE: 2001-03-21
 / PRIOR APPLICATION NUMBER: 60/191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-23
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR FILING DATE: 2000-12-22
 / NUMBER OF SEQ ID NOS: 14110
 / SOFTWARE: FastSEQ for Windows Version 4.0
 / SEQ ID NO: 13818
 / LENGTH: 940
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 / US-09-815-242-13818

Query Match 60.0%; Score 48; DB 9; Length 940;
 Best Local Similarity 64.3%; Pred. No. 1.5e+02;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 .LADLYBEGGGGGGB 15
 Db 901 IVDLGPEGGGGGB 914.

RESULT 15
 US-09-815-242-13818
 / Sequence 13818, Application US/09815242
 / Patent No. US2002006156A1
 / GENERAL INFORMATION:
 / APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Kari L.
 / APPLICANT: Zyskind, Judith W.
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John D.
 / APPLICANT: Carr, Grant J.
 / APPLICANT: Yamamoto, Robert T.
 / APPLICANT: Xu, H. Howard
 / TITLE OF INVENTION: Identification of Essential Genes in
 / Prokaryotes

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

ON Protein - protein search, using SW model

Run on: December 30, 2003; 12:00:03 ; Search time 21 Seconds
(without alignments)
68.692 Million cell updates/sec

Title: US-09-807-610-11
Perfect score: 80
Sequence: 1 ALADLYEGGGGGG 15
Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

4: Pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	80	100.0	180	2	A39386	interleukin-1 receptor antagonist, long intracellular splice form - human
2	49	61.3	206	2	T32854	interleukin-1 receptor antagonist, short intracellular splice form
3	48	60.0	325	2	G96718	N; Contains: interleukin-1 receptor antagonist, short intracellular splice form
4	48	60.0	940	1	BVECTA	C; Species: Homo sapiens (man)
5	48	60.0	940	2	H91258	C; Date: 28-Feb-1992 #sequence_revision 11-Apr-1997 #text_change 26-May-2000
6	48	60.0	940	2	D86099	R; Accession: I37893; A39386
7	48	60.0	941	2	AB1017	R; Muzio, M.; Polentanitti, N.; Sironi, M.; Poli, G.; De Gioia, L.; Introna, M.; Mantovan, J. Exp. Med. 182, 623-628, 1995
8	48	60.0	947	2	AP0040	J; Title: Cloning and characterization of a new isoform of the interleukin 1 receptor antagonist, long intracellular splice form of the human interleukin 1 receptor antagonist, short intracellular splice form
9	47	58.8	201	2	JQ1094	A; Reference number: A39386; MUTIN:9535865; PMID:7629520
10	47	58.8	726	2	G83310	A; Status: translated from GB/EMBL/DDJB
11	46.5	58.1	158	2	F96592	A; Molecule type: mRNA
12	46.5	58.1	455	2	B86427	A; Residues: 1-180 <REG>
13	46	57.5	269	2	A70119	A; Cross-references: EMBL:X04348; PIDN:9100970; PIDN:CAA5087; PID:9108971
14	46	57.5	311	2	F85141	C; Comment: For an alternative splice form, see PIR:A30368
15	46	57.5	385	2	T20110	C; Genetics:
16	46	57.5	578	2	T48195	A; Gene: GDB:IL1RN
17	46	57.5	581	2	T22341	A; Cross-references: GDB:125897; OMIM:147679
18	46	57.5	750	2	T10644	A; Map location: 2q14.2-2q14.2
19	45.5	56.9	431	1	BME10	C; Superfamily: interleukin-1
20	45.5	56.9	1171	2	F83110	C; Keywords: alternative splicing; cytokine receptor
21	45	56.2	386	1	S22315	F1-180/Product: interleukin-1 receptor antagonist, long intracellular splice form #stat1 F1-3,25-180/Product: interleukin-1 receptor antagonist, short intracellular splice form
22	45	56.2	396	2	S55820	Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 15; Conservativeness 0; Indels 0; Gaps 0;
23	45	56.2	404	2	S16468	QY 1 ALADLYEGGGGGGG 15
24	45	56.2	404	2	S55729	DB 2 ALADLYEGGGGGGG 16
25	45	56.2	629	2	F84248	DB 2 ALADLYEGGGGGGG 16
26	45	56.2	653	2	C86448	DB 2 ALADLYEGGGGGGG 16
27	45	56.2	856	2	G96814	DB 2 ALADLYEGGGGGGG 16
28	45	56.2	943	2	D61057	DB 2 ALADLYEGGGGGGG 16
29	45	56.2	956	2	AB1761	DB 2 ALADLYEGGGGGGG 16

RESULT 2
T22854
R; Description: Hypothetical protein Y8G1A.1 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C; Accession: T32854
R; Cordes, M.; Le, T. T.
submitted to the EMBL Data Library, December 1997
A; Description: The sequence of C. elegans cosmid Y8G1A.

RESULT 8
 excinuclease ABC chain A *uvrA* [imported] - *Yersinia pestis* (strain C092)
 C;Species: *Yersinia pestis*
 C;Accession: AP0040
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; den-Dier-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skellton, J.; Stevens, R.; Whitehead, S.; Barrell, B.; Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AP0040
 A;Molecule type: DNA
 A;Status: preliminary
 A;Residues: 1-940 <STO>
 A;Cross-references: GB:AB005174; NID:gi12519009; PIDN:AGC59256.1; GSPDB:GN00145; UWGZ:256
 A;Experimental source: strain Q57:17; substituin EPU933
 C;Genetics:
 A;Gene: *uvrA*
 C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 60.0%; Score 48; DB 2; Length 940;
 Best Local Similarity 64.3%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 LADLYTBGGGGGB 15
 Db 901 IVDLGPFGGGGG 914

RESULT 9
 AB1094
 excinuclease chain A [imported] - *Salmonella enterica* subsp. *enterica* serovar *Typhi*
 C;Species: *Salmonella enterica* subsp. *enterica* serovar *Typhi*
 A;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AB1017
 R;Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, N.; Moule, S.; O'Gara, P.
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar *Agona*
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AB1017
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-341 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD09238.1; PID:916505242; GSPDB:GN00176
 C;Genetics:
 A;Gene: *uvrA*
 C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 60.0%; Score 48; DB 2; Length 941;
 Best Local Similarity 64.3%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 LADLYTBGGGGGB 15
 Db 901 IVDLGPFGGGGG 914

RESULT 10
 G83310
 conserved hypothetical protein PA2685 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: G83310
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradihan, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406: 959-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A;Reference number: AB2950; MUID:20437337; PMID:10984043
 A;Accession: G83310
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-726 <STO>
 A;Cross-references: GB:AB004696; GB:AB004091; NID:99948750; PIDN:AA006073.1; GSPDB:GN001:
 C;Genetics:
 A;Gene: PA2685

RESULT 11
 F96692

hypothetical protein T4024.4 [Imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: P86692
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.; Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A.;Authors: Salberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-158 <STO>
 A;Cross-references: GB:AB005173; NID:g11128395; PIDN:ARG31200.1; GSPDB:GN00141
 C;Genetics:
 C;Gene: TAO24.4
 A;Map position: 1

Query Match Score 46.5%; DB 2; Length 158;
 Best Local Similarity 47.6%; Pred. No. 9.3;
 Matches 10; Conservative 2; Mismatches 2; Indels 7; Gaps 1;
 Qy 2 IADLY-----BEGGGGGG 15
 Db 72 IGDLYVSSPPPBEGGGGG 92

RESULT 12
 B86427 hypothetical protein F12P21.7 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C;Accession: B86427
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.; Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, A.;Authors: Salberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-155 <STO>
 A;Cross-references: GB:AB005172; NID:g11120781; PIDN:ARG30962.1; GSPDB:GN00141
 C;Genetics:
 C;Gene: TAO24.4
 A;Map position: 1

Query Match Score 46.5%; DB 2; Length 455;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 Qy 4 DLY-BEGGGGGG 14
 Db 295 DRYBEGGGGG 306

RESULT 13
 A70719 probable enoyl-CoA hydratase - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: A70719
 A;Status: preliminary; translated from GB/EMBL/DDBJ

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skeilton, S.; Squares, S.;
 A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Accession number: A70500; MUID:98295987; PMID:9631230
 A;Accession: A70719
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-269 <CDS>
 A;Cross-references: GB:279700; GB:AL123456; NID:93261628; PIDN:CAB02007.1; PMID:91524191
 A;Experimental source: strain H37RV
 C;Genetics:
 C;Gene: echa7
 C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology
 P;29-183?Domain: enoyl-CoA hydratase homology <ECH>

Query Match Score 46%; DB 2; Length 269;
 Best Local Similarity 81.8%; Pred. No. 19;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 ADLVBEGGGCG 13
 Db 68 ADLSEAGGGGG 78

RESULT 14
 B8641 hypothetical protein F9R16.12 - Arabidopsis thaliana (mouse-ear cress)
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C;Accession: F86341
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.; Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, S.L.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, A.;Authors: Salberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: F86341
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-311 <STD>
 A;Cross-references: GB:AB005172; NID:94836899; PIDN:AAD30602.1; GSPDB:GN00141
 C;Genetics:
 C;Gene: TAO24.4
 A;Map position: 1
 C;Superfamily: Arabidopsis thaliana hypothetical protein T12H17.200

Query Match Score 46%; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 BEGGGGGG 14
 Db 249 BEGGGGGG 256

RESULT 15
 T20410 hypothetical protein B02A10.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T20410
 R;Thomas, K.
 Submitted to the EMBL Data Library, October 1996
 A;Reference number: Z92721
 A;Accession: T20410
 A;Status: preliminary; translated from GB/EMBL/DDBJ

A; Molecule type: DNA
A; Residues: 1-385 <WIL>
A; Cross-references: ENBL:281053; PION:CAB02877.1; GSPDB:GN00023; CESP:E02A10.2
A; Experimental source: clone E02A10
C; Genetics:
A; Gene: CESP:E02A10.2
A; Map position: 5
A; Introns: 32/1; 72/1; 85/1; 122/1; 133/1; 220/3

Query Match 57.5%; Score 46; DB 2; Length 385;
Best Local Similarity 64.3%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 4; Indels 0;
Gaps 0;

Qy 1 ALADIVSEGCGGG 14
 | : |||||
 15 AT AFLPPSGGGGG 28

Search completed: December 30, 2003, 12:03:20
Job time : 22 secs

Result No.	Score	Query Match Length	DB ID	Description
1	50	62.5	977 1 Y848_HUMAN	09433 homosapien
2	48	60.0	576 1 2384_HUMAN	Q8t68 homosapien
3	48	60.0	579 1 2384_RAT	O9ej4 rattus norvegicus
4	48	60.0	940 1 UVRA_BCO57	Q8xs9 escherichia coli
5	48	60.0	940 1 UVRA_ECOL6	Q8fb02 escherichia coli
6	48	60.0	940 1 UVRA_ECOLI	P07671 escherichia coli
7	48	60.0	941 1 UVRA_SALTY	P37434 salmonella enterica
8	48	60.0	947 1 UVRA_TERPB	Q8zj07 yersinia pestis
9	47	58.8	201 1 YR21_TRSVR	P25245 tomato ring virus
10	46	57.5	311 1 EBCA_ARATH	Q9s7c9 arabidopsis thaliana
11	46	57.5	1175 1 HCN4_RABIT	Q9tv66 orctolagus cuniculus
12	45.5	57.9	431 1 BMP7_HUMAN	P18075 homo sapiens
13	45	56.2	386 1 RB87_DROME	P48810 drosophila melanogaster
14	45	56.2	396 1 YJB8_BEAST	P47049 saccharomyces cerevisiae
15	45	56.2	404 1 C4Z_DROME	Q27294 drosophila melanogaster
16	45	56.2	404 1 SCJ1_YEAST	P25303 saccharomyces cerevisiae
17	45	56.2	490 1 JIP1_DROMB	Q9wko0 drosophila melanogaster
18	45	56.2	943 1 UVRA_HABIN	P44410 haemophilus ducreyi
19	45	56.2	973 1 UVRA_RHILLO	Q98m36 rhizobium leguminosarum
20	44	55.0	199 1 AC22_TBNSMO	P26168 tenebrio molitor
21	44	55.0	342 1 ROA1_SCHEM	P21522 schistosoma mansoni
22	44	55.0	358 1 CEBA_HUMAN	P49115 homo sapiens
23	44	55.0	377 1 GP27_RAT	Q9ijh3 rattus norvegicus
24	44	55.0	601 1 HR78_DROME	Q24142 drosophila melanogaster
25	44	55.0	943 1 UVRA_PASMO	P57797 pastorella pastorelli
26	43	53.8	305 1 ROA1_HUMAN	Q13351 homo sapiens
27	43	53.8	375 1 GP27_HUMAN	Q9hs67 homo sapiens
28	43	53.8	460 1 SR54_HALN1	Q9mans halobacteri
29	43	53.8	502 1 XYLB_LACBR	P35850 lactobacillus
30	43	53.8	515 1 PVR1_MOUSE	Q91kf6 mus musculus
31	43	53.8	517 1 PVR1_HUMAN	Q15223 homo sapiens
32	43	53.8	642 1 HS72_PICAN	P53623 pichia angusta
33	43	53.8	687 1 WRK2_ARATH	Q9f977 arabidopsis thaliana

PT	DOMAIN	712 AA;	723	POLY-GLY.
SQ	SEQUENCE	977 MW;	3C936B7E003DF54	CRC64;
Query Match		62.5%;	Score: 50;	DB 1; Length 977;
Best Local Similarity		80.0%;	Pred. No. 12;	
Matches		8;	Mismatches	0;
			Indels	0;
Gaps				0;
Qy	5 LYBEGGGGG 14			
	: :			
Db	708 LFEDGGGGG 717			
RESULT 2				
Z384_HUMAN	STANDARD;	PRT;	576 AA.	
ID	Z384_HUMAN	Q8TPF8; O15407;		
AC		Q8NP38;		
DT	28-PFB-2003	(Rel. 41; Created)		
DT	28-PFB-2003	(Rel. 41; Last sequence update)		
DT	15-SSP-2003	(Rel. 42; Last annotation update)		
DB	Zinc finger protein 384	(Nuclear matrix transcription factor 4)		
DB	(CAG repeat protein 1).			
GN	ZNF384	OR NM_01204 OR CAGHI.		
OS	Homo sapiens (Human).			
OC	Eukaryota; Metzoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RA	Matsumo M. Y.,			
RL	Submitted. (AUG-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Brain;			
RA	Tanigami A., Fujiiwara T., Shishibori T., Goto Y., Hirao M., Shimizu F.,			
RA	Wakabe H., Ono T., Hischiki H., Watanabe T., Ozaki K., Sugiyama T.,			
RA	Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,			
RA	Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,			
RA	Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,			
RA	Wagatsuma M., Murakawa K., Kanehori K., Takahashi-T Fujii A., Oshima A.,			
RA	Sugiyama B., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,			
RA	Masuhisa Y., Nagai K., Isegai T.,			
RT	"NEDO human cDNA sequencing project."			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 395-576 FROM N.A.			
RC	TISSUE=Brain cortex			
RA	Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kidwell A.S.,			
RA	Breschel T.S., Stine C., Callahan C., McInnis M.G., Ross C.A.;			
RT	"cDNAs with long CAG trinucleotide repeats from human brain."			
RL	Hum. Genet. 100:114-122(1997).			
CC	- FUNCTION: Transcription factor that binds the consensus DNA			
CC	sequence (GC)AAAATA. Seems to bind and regulate the promoters of			
CC	MMP1, MMP3, MMP7 and COL1A1 (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (By similarity).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=2;			
CC	Comment=Additional isoforms seem to exist;			
CC	Name=;			
CC	Isold=Q8TPF68-1; Sequence=Displayed;			
CC	Isold=Q8TPF68-2; Sequence=VSP 006920;			
CC	-1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-PINGER PROTEINS.			
CC	-1- SIMILARITY: Contains 8 C2H2-type zinc fingers.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)			
CC	or send an email to license@isb-sib.ch .			
CC	CC	CC	CC	CC
DR	EMBL; AB070238; BAB85125.1;			
DR	EMBL; AK095734; BAC04618.1;			
DR	EMBL; U80738; AAB91437.1;			
DR	Genew; HGNC:11955; ZNP784.			
DR	InterPro; IPR007087; Znf_C2H2.			
DR	Pfam; PF00096; zf-C2H2; 8.			
DR	ProDom; PD000003; Znf_C2H2; 4.			
DR	SMART; SM00355; Znf_C2H2; 8.			
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.			
DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.			
KW	Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;			
KW	DNA-binding; Regulation; Zinc-finger; Metal-binding; Nuclear protein;			
PT	ZN_FINGER	228	C2H2-TYPE 1.	
PT	ZN_FINGER	250	C2H2-TYPE 2.	
PT	ZN_FINGER	256	C2H2-TYPE 2.	
PT	ZN_FINGER	278	C2H2-TYPE 3.	
PT	ZN_FINGER	306	C2H2-TYPE 3.	
PT	ZN_FINGER	317	C2H2-TYPE 4.	
PT	ZN_FINGER	345	C2H2-TYPE 5.	
PT	ZN_FINGER	373	C2H2-TYPE 6.	
PT	ZN_FINGER	403	C2H2-TYPE 7.	
PT	ZN_FINGER	433	C2H2-TYPE 8.	
PT	DOMAIN	461	GLN-RICH.	
PT	DOMAIN	466	ALA-RICH.	
PT	VARSPLIC	300	Missing (in isoform 2).	
PT	VARSPLIC	360	/FtId=VSP 006920.	
SEQ	SEQUENCER	576 AA;	63091 MW;	2A152786C3C16D0 CRC64;
Query Match		60.0%;	Score: 48;	DB 1;
Best Local Similarity		90.0%;	Length 576;	
Matches		9;	Pred. No. 13;	
			Mismatches 0;	
			Indels 0;	
			GapS 0;	
Qy	5 LYBEGGGGG 14			
TD	2384 RAT	STANDARD;	PRT;	579 AA.
AC	Q8TPQ4; Q9BQJ3; Q9BQJ5;			
DT	Q8TPQ4; Q9BQJ2; Q9BQJ3; Q9BQJ5;			
DT	28-FEB-2003 (Rel. 41; Created)			
DT	28-FEB-2003 (Rel. 41; Last sequence update)			
DT	15-SEP-2003 (Rel. 42; Last annotation update)			
DB	172 LYBEGGGGG 181			
RESULT 3				
TD	2384 RAT	STANDARD;	PRT;	579 AA.
AC	Q8TPQ4; Q9BQJ2; Q9BQJ3; Q9BQJ5;			
DT	28-FEB-2003 (Rel. 41; Created)			
DT	15-SEP-2003 (Rel. 42; Last annotation update)			
DB	Zinc finger protein 384 (Nuclear matrix transcription factor 4)			
GN	(Cas-associated zinc finger protein).			
GN	ZNF384 OR NM_01204 OR CIZ.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Buthidae; Rodentia; Sciurognathini; Murinae; Rattus.			
OX	NCBITaxonID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, DNA-BINDING, AND			
RP	INTERACTION WITH CAS.			
RP	Subcellular location with CAS.			
RX	Thunyakirtipaisal P., Alvarez M., Tokunaga K., Onyia J.B., Hock J., Ohashi N., Feister H., Rhodes S.J., Bidwell J.P., Nakamoto T., Yamagata T., Sakai R., Honda H., Ueno H., Hirano N., Yazaki Y., Hirai H., Buthidae; Rodentia; Sciurognathini; Murinae; Rattus.			
RT	"CIZ, a zinc finger protein that interacts with p130cas and activates the expression of matrix metalloproteinases."			
RT	Mol. Cell. Biol. 20:1649-1658 (2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).			
RC	STRAIN=Prague-Dawley;			
RC	MEDLINE=21024193; PubMed=11149472;			
RA	Thunyakirtipaisal P., Alvarez M., Tokunaga K., Onyia J.B., Hock J., Ohashi N., Feister H., Rhodes S.J., Bidwell J.P., Nakamoto T., Yamagata T., Sakai R., Honda H., Ueno H., Hirano N., Yazaki Y., Hirai H., Buthidae; Rodentia; Sciurognathini; Murinae; Rattus.			
RT	"Cloning and functional analysis of a family of nuclear matrix transcription factors (NP/NMP4) that regulate type I collagen expression in osteoblasts."			
RT	J. Bone Miner. Res. 16:10-23 (2001).			
RL	-1- FUNCTION: Transcription factor that binds the consensus DNA sequence [GC]AAAAT. Seems to bind and regulate the promoters of			
CC	MMMP1, MMMP3, MMP7 and COL1A1.			
CC	MMMP1, MMMP3, MMP7 and COL1A1.			

CC	- - SUBUNIT: Interacts with Cas.
CC	- - SUBCELLULAR LOCATION: Nuclear.
CC	- - ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=3;
CC	Comment=additional isoforms seem to exist;
CC	Name=1;
CC	IsoId=Q9RQJ4-1; Sequence=Displayed;
CC	Name=2;
CC	IsoId=Q9RQJ4-2; Sequence=VSP_006921;
CC	Names=3;
CC	IsoId=Q9RQJ4-3; Sequence=VSP_006922;
CC	- - TISSUE SPECIFICITY: Expressed in osteocytes, osteoblasts, and chondrocytes in bone.
CC	- - SIMILARITY: BELONGS TO THE KRUPPBL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.
CC	- - SIMILARITY: Contains 8 C2H2-type zinc fingers.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; AB019281; BAA89667; 1.
CC	EMBL; AP216804; AAG40582; 1.
CC	EMBL; AP216805; AAG40583; 1.
CC	EMBL; AP216806; AAG40584; 1.
CC	HSSP; P08153; 1 ZFP.
CC	DR TRANSFAC; T05136; .
CC	DR TRANSFAC; T05137; .
CC	DR TRANSFAC; T05138; .
CC	DR TRANSFAC; T05141; .
CC	DR TRANSFAC; T05142; .
CC	InterPro; IPR007087; Znf_C2H2.
CC	Pfam; PF00096; zf-C2H2; 6.
CC	ProDom; PD000003; zf_C2H2; 4.
CC	SMART; SM00355; Znf_C2H2; 8.
CC	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
CC	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.
KW	Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein; DNA binding; Repeat; Alternative splicing.
KW	ZN FING 229 251 C2H2 TYPE 1.
KW	FT 2Z 257 279 C2H2 TYPE 2.
KW	FT 2Z 285 307 C2H2 TYPE 3.
KW	FT 2Z 318 340 C2H2 TYPE 4.
KW	FT 2Z 346 368 C2H2 TYPE 5.
KW	FT 2Z 374 398 C2H2 TYPE 6.
KW	FT 2Z 404 426 C2H2 TYPE 7.
KW	FT 2Z 434 456 C2H2 TYPE 8.
KW	DOMAIN 462 524 GLN-RICH.
KW	VARSPLIC 467 506 ALA-RICH.
KW	VARSPLIC 103 118 Missing (in Isoform 2).
KW	VARSPLIC 103 118 /Pfam=VSP_006921.
KW	VARSPLIC 301 361 Missing (in Isoform 3).
KW	VARSPLIC 301 361 /Pfam=VSP_006922.
KW	CONFLICT 178 179 GG -> RS (IN REF. 1).
KW	CONFLICT 576 577 LA -> WP (IN REF. 1).
SEQ	SEQUENCE 579 AA; 63139 MW; PBC24280D1050C45 CRC64;
Query Match	60.0%
Best Local Similarity	90.0%
Matches	9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DY	5 LYBEGGGGGG 14
Dbb	172 LTBEGGGGG 181
RESULT 4	
ID	JYRA_ECO57
ID	UVRA_ECO57 STANDARD; PRT; 940 AA.

Best Local Similarity 64.3%; Pred. No. 21; Mismatches 1; Indels 4; Gaps 0; Qy 2 LADLYBEGGGGG 15 Db 901 IVDLGPGGSSGGB 914

RESULT 5

UVRA_ECOLI6 STANDARD; PRT; 940 AA.

ID UVRA_ECOLI6

AC QPB057;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DB UvraABC system protein A (Uvra protein) (Excinuclease ABC subunit A).

GN UVRA OR C5048.

OS Escherichia coli O6.

OC Enterobacteriaceae; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID217992;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN06 H1 / CFT073 / ATCC 700928;

RX MEDLINE:22471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Rasko D., Buckles B.L., Liou S.-R., Boutin A., Hackert J., Stroud D.,

RA Mayhew H.J.T., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.,

RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli".

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

CC -1. FUNCTION: The UvraBC repair system catalyzes the recognition and processing of DNA lesions. Uvra is an ApnApe and a DNA-binding protein. A damage recognition complex composed of 2 Uvra and 2 UvRB subunits scans DNA for abnormalities. When the presence of a lesion has been verified by UvRB, the UvRA molecules dissociate

CC ^ (By similarity)

CC -1 SUBUNIT: Forms a heterotetramer with UvRB during the search for lesions. (By similarity).

CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AE016770; AAN03474.1; -.

DR HAMAP; MF_00205; -; 1.

DR InterPro; IPR003439; ABC transporter.

DR InterPro; IPR004602; Uvra.

DR Pfam; PF00005; ABC transporter.

DR ProDom; PD000006; ABC transporter; 1.

DR TIGRFAMS; TIGR00630; Uvra; 1.

DR PROSITE; PS00211; ABC TRANSPORTER_1; 2.

DR PROSITE; PS000831; ABC_TRANSPORTER_2; 1.

KW SOS response; Excision nuclease; DNA repair; DNA recombination;

KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;

KW Zinc-finger; Complete proteome.

PT NP_BIND 31 38 ATP (POTENTIAL).

PT ZN_FING 253 280 C4-TYPE.

PT NP_BIND 640 647 ATP (POTENTIAL).

PT ZN_FING 740 766 C4-TYPE.

SQ SEQUENCE 940 AA; 103882 MN; A20C90C53816ACEB CRC64;

Qy 2 LADLYBEGGGGG 15

Db 901 IVDLGPGGSSGGB 914

RESULT 6

UVRA_ECOLI STANDARD; PRT; 940 AA.

ID UVRA_ECOLI

AC P07671; P76788;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1998 (Rel. 07, Last sequence update)

DT 28-FBB-2003 (Rel. 41, Last annotation update)

GN UVRA OR DINE OR B4058 OR SF4146.

OS Escherichia coli, and

OS Shigella flexneri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Bacteroidales; Escherichia.

NCBI_TaxID=562, 623;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli;

RX MEDLINE:86168204; PubMed=3007478;

RA Husain I., van Houten B., Thomas D.C., Sancar A.;

RT "Sequences of Escherichia coli uvra gene and protein reveal two potential ATP binding sites.";

RL J. Biol. Chem. 261:4895-4901(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12 / MG1655;

RX MEDLINE:94089392; PubMed=8265357;

RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.;

RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes.";

RL Nucleic Acids Res. 21:5408-5417(1993).

RN [3]

RP SEQUENCE OF 1-25 FROM N.A.

RC SPECIES=E.coli;

RX MEDLINE:83292521; PubMed=6110514;

RA Backendorf C., Brandts J.A., Kartasova T., van de Putte P.;

RT "In vivo regulation of the uvra gene: role of the -10, and -35' promoter regions";

RL Nucleic Acids Res. 11:5795-5810(1983).

RN [4]

RP SEQUENCE OF 1-14 FROM N.A.

RC SPECIES=E.coli;

RX MEDLINE:82220077; PubMed=6483374;

RA Sancar A., Sancar G.B., Rupp W.D., Little J.W., Mount D.W.;

RT "LexA protein inhibits transcription of the E. coli uvra gene in vitro.";

RT Nature 298:96-98(1982).

RN [5]

RP SEQUENCE FROM N.A.

RC SPECIES=S.flexnei; STRAIN=301 / Serotype 2a;

RX MEDLINE:222272406; PubMed=12384530;

RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Yang J., Yang F., Zhang J., Yang G., Wu H., Ou D., Dong J., Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S., Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y., Yu J.;

RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157.";

RT Nucleic Acids Res. 30:4432-4441(2002).

RN [6]

RP CHARACTERIZATION

RC SPECIES=E.coli;

RX MEDLINE:91208117; PubMed=1826851;

RA Myles G.M., Sancar A.;

RT "Isolation and characterization of functional domains of UvraA.";

RL Biochemistry 30:3834-3840(1991).

RN [7]

RP MUTAGENESIS OF CYS-253.

RC SPECIES=B.coli;
 RX MEDLINE=89380205; Published=2250431;
 RA Navaratnam S., Myles G.M., Strange R.W., Sancar A.;
 RT "Evidence from extended X-ray absorption fine structure and site-
 specific mutagenesis for zinc fingers in UvrA protein of Escherichia
 coli."
 RT J. Biol. Chem. 264:16067-16071(1989).
 CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and
 processing of DNA lesions. UvrA is an ATPase and a DNA-binding
 protein. A damage recognition complex composed of 2 uvrA and 2
 uvrB subunits scans DNA for abnormalities. When the presence of a
 lesion has been verified by uvrB, the uvrA molecules dissociate
 (By similarity).
 CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
 12-basepair (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- MISCELLANEOUS: Binds about 2 zinc atoms/molecule.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR A SUBFAMILY.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC DR EMBL; M13495; AAA24754.1; -.
 CC DR EMBL; U00066; AAC43152.1; -.
 CC DR EMBL; AE000479; AAC7028.1; -.
 CC DR EMBL; X01621; CAA25764.1; -.
 CC DR EMBL; J01721; AAA24753.1; -.
 CC DR EMBL; AE015423; AACN5568.1; -.
 CC PIR; A23861; BVECA.
 CC DR ECO2DBASE; H1124.0; 6TH EDITION.
 CC DR Ecogene; BG11061; uvrA.
 CC DR HAMAP; MF_00205; -; 1.
 CC DR InterPro; IPR003439; ABC transporter.
 CC DR InterPro; IPR004602; UvrA.
 CC DR Pfam; PF00005; ABC_tran; 1.
 CC DR ProDom; P000006; ABC transporter; 1.
 CC DR TIGRFAMS; TIGR00630; uvrA; 1.
 CC DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 CC DR PROSITE; PS50083; ABC_TRANSPORTER_2; 1.
 CC KW DNA excision; ATP-binding; DNA repair; DNA recombination;
 CC KW Zinc-finger; Complete proteome.
 CC FT NP_BIND 31 38 ATP.
 CC FT ZN_FING 253 280 CA4-TYPE.
 CC FT NP_BIND 640 647 ATP.
 CC FT ZN_FING 740 766 CA4-TYPE.
 CC FT MUTAGEN 253 253 C->A,H,S: REDUCED ACTIVITY.
 SQ SEQUENCE 940 AA; 103867 MN; D61AAEB6514BB60 CRC64.

Query Match Score 60.0%; Best Local Similarity 64.3%; Pid. No. 21; Length 940; Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IADLYTBEGGGGGG 15
 Db 901 IVDLGPEGGGGG 914

RESULT 7
 UVRA_SALTY STANDARD PRT; 941 AA.
 AC P3734;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SBP-2003 (Rel. 42, Last annotation update)
 DB UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
 GN UVRA OR STM4234 OR STM4450 OR T4160.
 JS Salmonella typhimurium, and

PFam: PF00005; ABC tran; 2.
DR ProdDom; PD00006; ABC transporter; 1.
DR TIGRFAMS; TIGR00630; uvra; 1.
DR PROSITE; PS00211; ABC TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC TRANSPORTER_2; 1.
KW DNA response; Excision nucleases; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.

PT NP BIND 31 38 ATP.
PT ZN FING 253 280 CA-TYPE.
PT NP-BIND 640 647 ATP.
PT ZN FING 740 766 C4-TYPE.
SQ SEQUENCE 941 AA; 103928 MW; CAAFC9P549060C26 CRC64;

Query Match 60.0%; Score 48; DB 1; Length 941;
Best Local Similarity 64.3%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LADLYBERGGGGGB 15
Db 901 YVDLGPEGGSGGGB 914

RESULT 8
UVRA_YERP
ID UVRA_YERP STANDARD PRT; 947 AA.

AC Q82J07; DT 15-SEP-2003 (Rel. 42; Created)
DT 15-SEP-2003 (Rel. 42; Last sequence update)

DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE UvRABC system protein A (UvRA protein) (Excinuclease ABC subunit A).
GN UVRA OR Ypo0324 OR Y0580.

OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

NCBI_Taxid632; RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Co-92 / Biovar Orientalis;
RX MEDLINE=20470413; PubMed=11583160;

RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

RA Prentice M.B., Sebastian M.A., James K.D., Churcher C., Mungall K.L.,

RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Feltwell T., Hamlin N., Holroyd S., Jages S., Karlyshev A.V.,

RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutledge K.,

RA Simmonds M., Shelton J., Stevens K., Whitehead S., Barrell B.G.;

RT Genome sequence of *Yersinia pestis*, the causative agent of plague.;
RN Nature 413:523-527(2001).

[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III., Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

RA Peterston J.D., Lindler L.B., Brubaker R.R., Piano G.V.,

RA Sperry K.A., McDonough K.A., Nilles M.R., Matson J.S., Blattner F.R.,

RA Perry R.D.;

RT "Genome sequence of *Yersinia pestis* KIM.";
RL Bacteriol. 184:4601-4611(2002).

CC FUNCTION: The UvRABC repair system catalyzes the recognition and
processing of DNA lesions. Uvra is an APnase and a DNA-binding
protein. A damage recognition complex composed of 2 uvrA and 2
uvrB subunits scans DNA for abnormalities. When the presence of a
lesion has been verified by uvrB, the uvrA molecules dissociate
(BY SIMILARITY).
CC -I- SUBUNIT: Forms a heterotetramer with uvrB during the search for
lesions (BY SIMILARITY).

CC -I- SUBUNIT LOCATION: Cytoplasmic (BY SIMILARITY).

CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC or send an email to license@ib-sib.ch/announce/
CC EMBL; AJ114142; CAC89185..1;
CC EMBL; AB013660; AAC84168..1;
CC HAMAP; MF_00205; -; 1;
CC InterPro; IPR003433; ABC transporter.
CC InterPro; IPR004602; UvRA.
CC Pfam; PF00005; ABC tran, 1.
CC ProDom; PD00006; ABC transporter; 1.
CC TIGRFAMS; TIGR00630; uvrA; 1.
CC PROSITE; PS00211; ABC TRANSPORTER_1; 2.
CC PROSITE; PS5093; ABC TRANSPORTER_2; 1.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.

FT NP BIND 31 38 ATP.
FT ZN FING 253 280 CA-TYPE.
FT NP-BIND 640 647 ATP.
FT ZN FING 740 766 C4-TYPE.
SQ SEQUENCE 941 AA; 103928 MW; CAAFC9P549060C26 CRC64;

Query Match 60.0%; Score 48; DB 1; Length 947;
Best Local Similarity 64.3%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LADLYBERGGGGGB 15
Db 901 YVDLGPEGGSGGGB 914

RESULT 9
YR21_TRSVR STANDARD PRT; 201 AA.

ID YR21_TRSVR STANDARD PRT; 201 AA.
AC P25145; DT 01-MAY-1992 (Rel. 22; Created)

DT 01-MAY-1992 (Rel. 22; Last sequence update)

DT 15-SEP-2003 (Rel. 42; Last annotation update)

DB Hypothetical 20.2 kDa protein in RNA2.

OS Tomato ringspot virus (isolate raspberry) (TomRSV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;

OC Nepoviridae.

NCBI_TaxID=12281; RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=9131402; PubMed=1856689;

RA Rott M.B., Tremaine J.H., Rochon D.M.; RT "Nucleotide sequence of tomato ringspot virus RNA-2.";

RT RL Virol. 72:1505-1514(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ib-sib.ch/announce/>
CC or send an email to license@ib-sib.ch).

CC EMBL; DI2477; BA02044..1;
CC DR PIR; J01094; J01094.

CC KW HYPOTHETICAL PROTEIN.

CC FT DOMAIN 15 22 POLY-GLY.

CC PT DOMAIN 61 66 POLY-GLY.

CC SQ SEQUENCE 201 AA; 20194 MW;

Query Match 58.8%; Score 47; DB 1; Length 201;
Best Local Similarity 75.0%; Pred. No. 6;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy	RESULT 10	ESCA_ARATH	STANDARD;	PRT;	311 AA.	0;
Db	Q9STC9;					
	DT 28-FEB-2003 (Rel. 41, Created)					
	DT 28-FEB-2003 (Rel. 41, Last sequence update)					
	DT 28-FEB-2003 (Rel. 41, Last annotation update)					
	Putative DNA-binding protein ESCAROLA.					
	DB ESS_OG_ARIG2090 OR P9H16.12.					
	EN					
	DN					
	DC Arabidopsis thaliana (Mouse-ear cress)					
	DC Sporophytidae; Magnoliophyta; Streptophyta; Embryophyta; Tracheophyta; core eudicots; Rosidae;					
	DC eurosid II; Brassicales; Brassicaceae; Arabidopsis.					
	DX NCBI_TAXID=3702;					
	RN					
	SEQUENCE FROM N.A.					
	SPRAIN=cv_Columbia; PubMed=10759495;					
	RC MEDLINE=20223715;					
	RX Weigel D., Ahn J.H., Blazquez M.A., Borevitz J.O., Christensen S.K.,					
	RX Fankhauser C., Ferrandiz C., Kardailsky I., Malanchukovil B.J.,					
	RX Nef M.M., Nguyen J.T., Sato S., Wang Z.Y., Xia Y., Dixon R.A.,					
	RX Harrison M.J., Lamb C.J., Yanofsky M.P., Chory J.,					
	RX "Activation tagging in Arabidopsis." ;					
	RX Plant Physiol. 122:1003-1013(2000).					
	RN [1]					
	SEQUENCE FROM N.A.					
	SP STRAIN=Columbia; PubMed=11130712;					
	RX MEDLINE=21016719;					
	RX Theologis A., Ecker J.R., Palm C.J., Pederspiel N.A., Kaul S.,					
	RX Alonso O., Araujo R., Bowman C.L., Brooks S.Y.,					
	RX Blehler B., Chan A., Chao O., Chen H., Chenk R.F., Chin C.W.,					
	RX Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,					
	RX Dunn P., Etgu P., Feldblum T.Y., Feng J.-D., Fong B., Fujii C.Y.,					
	RX Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,					
	RX Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,					
	RX Kim C.J., Koo H.L., Kremenetskaya I., Kurtz J., Lam B.,					
	RX Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.,					
	RX Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,					
	RX Militscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,					
	RX Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,					
	RX Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,					
	RX Sun H., Talon L.J., Tambang G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,					
	RX Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,					
	RX "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana." ;					
	RX Nature 408:816-820(2000).					
	RN [2]					
	CC -1- MISCELLANEOUS: Overexpression causes late flowering and modified leaf development.					
	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
	CC EMBL; AP194974; AAP07197.1; -.					
	DR EMBL; AC007369; AAD30602.1; -.					
	DR PIR; F86341; F86341; -.					
	DR InterPro; IPR00537; AT_hook.					
	DR InterPro; IPR005175; DUP296.					
	DR Pfam; PF03479; DUP296; 1.					
	DR SMART; SM00384; AT_hook; 1.					
	CC DNA-binding.					
	CC DOMAIN BIND	86	98			A T HOOK (POTENTIAL)
	CC DOMAIN	26	52			GLY-RICH.

- DR InterPro; IPR00595; cnpB binding.
 DR InterPro; IPR00582; Ion_trans.
 DR InterPro; IPR005820; McFannell_nlq.
 DR Pfam; PF00520; Ion_trans; 1.
 SMART; SM00100; cNIP; 1.
 PROSITE; PS00888; CNMP_BINDING_1; 1.
 PROSITE; PS00889; CNMP_BINDING_2; FALSE_NEG.
 DR PROSITE; PS00842; CNMP_BINDING_3; 1.
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
 Potassium channel; Potassium; Potassium transport; Sodium transport;
 KW CAMP; cAMP-binding; Transmembrane; Glycoprotein.
 PT DOMAIN 1 267 CYTOPLASMIC (POTENTIAL).
 PT TRANSMEM 269 288 SEGMENT S1 (POTENTIAL).
 PT TRANSMEM 295 315 SEGMENT S2 (POTENTIAL).
 PT DOMAIN 316 341 CYTOPLASMIC (POTENTIAL).
 PT TRANSMEM 342 362 SEGMENT S3 (POTENTIAL).
 PT TRANSMEM 370 390 SEGMENT S4 (POTENTIAL).
 PT DOMAIN 391 421 CYTOPLASMIC (POTENTIAL).
 PT TRANSMEM 422 442 SEGMENT S5 (POTENTIAL).
 PT TRANSMEM 466 487 SEGMENT HS (PORE-FORMING) (POTENTIAL).
 PT TRANSMEM 498 518 SEGMENT S6 (POTENTIAL).
 PT DOMAIN 519 1175 CYTOPLASMIC (POTENTIAL).
 PT DOMAIN 210 261 INVOLVED IN SUBUNIT ASSEMBLY (BY
 SIMILARITY).
 PT DOMAIN 68 127 GLY-RICH.
 PT NP_BIND 596 711 CAMP.
 PT CARBOHYD 796 1160 PRO-RICH.
 PT SEQUENCE 459 459 N-LINKED (GLCNAC-). (POTENTIAL).
 SQ SEQUENCE 1175 AA; 126141 MW; 35A75PA9C710BD69 CRC64;
- Query Match Score 46; DB 1; Length 1175;
 Best Focal Similarity 64.3%; Pred. No. 51;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
- Qy 1 ALADLYBEGGGGG 14
 Db 4 :||| | ||| |
 4 105 SLASLGSRGCGGG 118
- RESULT 12 HUMAN STANDARD; PRT; 431 AA.
 ID BMP7_HUMAN P18075; Q9H512; Q9NTQ7;
 AC [1] _
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DB Bone morphogenetic protein 7 precursor (BMP-7) (Osteogenic protein 1).
 DS BMP7 OR OP1.
 DS Homo sapiens (Human)
 DC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 DC Bivalvia; Bivalvia; Primates; Catarrhini; Homidae; Homo.
- RN NCBI_TaxId=9606;
 RN [1] _
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=9291911; PubMed=2357959;
 RA Ozkaynak B.; Rueger D.C.; Drier B.A.; Corbett C.; Ridge R.J.;
 RA Sampath T.K.; Oppermann H.;
 RT "OP-1 cDNA encodes an osteogenic protein in the TGF-beta family."
 RL EMBO J. 9:12085-2093 (1990).
 RN [2]
 RP SEQUENCE FROM N.A. MEDLINE=91088608; PubMed=2261636;
 RA Celeste A.J.; Iannuzzi J.A.; Taylor R.C.; Hewick R.M.; Rosen V.;
 RA Wang E.A.; Wozney J.M.;
 RT "Identification of transforming growth factor beta family members
 present in bone-inductive protein purified from bovine bone."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9843-9847(1990).
 RN [3]
 RP SEQUENCE FROM N.A. MEDLINE=21638749; PubMed=11760052;
 RX MEDLINE=21638749; PubMed=11760052;
- RA Deloukas P.; Matthews L.H.; Burton J.; Ashurst J.; Gilbert J.G.R.;
 RA Jones M.; Stravrides G.; Almeida J.P.; Babbage A.K.; Baguley C.L.;
 RA Bailey J.; Barlow K.P.; Bates K.N.; Beard L.M.; Beare D.M.;
 RA Besley O.P.; Bird C.P.; Blahey S.B.; Bridgeman A.M.; Brown A.J.;
 RA Buck D.; Burrill W.D.; Butler A.P.; Carter C.; Carter N.P.; Cle C.M.;
 RA Chapman J.C.; Clamp M.; Clark G.; Clark L.N.; Clark S.Y.; Clee C.M.;
 RA Clegg S.; Cobley V.E.; Collier G.J.; Deadman R.; Dhami P.D.; Dunn M.;
 RA Coulson A.; Coville G.J.; Deadman R.; Connor R.E.; Corby N.R.;
 RA Ellington A.G.; Frankland J.A.; Fraser A.; French L.; Garner P.;
 RA Graham D.V.; Griffiths M.N.D.; Gwilliam R.; Hall R.E.;
 RA Hammond S.; Harley J.J.; Heath P.D.; Ho S.; Holden J.J.; Howden P.J.;
 RA Huckle B.; Hunt A.R.; Hunt S.B.; Jekosch K.; Johnson C.M.; Johnson D.;
 RA Kay M.P.; Kimberley A.M.; King A.; Knights A.; Laird G.K.; Lawlor S.;
 RA Lehyaesihaio M.H.; Leyvraz S.; Lovell J.D.; Lovell D.M.; Lovell J.D.;
 RA Marsh V.L.; Martin S.J.; McConnachie L.J.; McIay K.; McMurray A.A.;
 RA Milne S.A.; Misty D.; Moore M.J.F.; Mullikin J.C.; Nickerson T.;
 RA Oliver K.; Parker A.; Patel R.; Pearce T.A.V.; Peck A.I.;
 RA Philimore B.J.C.T.; Prathalingam S.R.; Plumb R.W.; Ramay H.;
 RA Shownkeen R.; Sirra K.; Steward C.A.; Sulston J.E.;
 RA Stice C.D.; Smith M.L.; Soderlund C.;
 RA Swann R.M.; Sycamore N.; Taylor R.; Tee L.; Thomas D.W.; Thorpe A.;
 RA Tracey A.; Trommsdorff A.C.; Vaudin M.; Wall M.; Wallis J.M.; Williams S.A.;
 RA Whitehead S.L.; Whittaker P.; Willey D.L.; Williams L.; Williams R.M.;
 RA Wilming L.; Wray P.W.; Hubbard T.; Durbin R.M.; Bentley D.R.; Beck S.;
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20."
 RL Nature 414:665-671 (2001).
- RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
- RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
 RA Klausner R.D.; Collins P.S.; Wagner L.; Schuler G.D.;
 RA Atkins C.M.; Ross M.T.; Scott C.B.; Sehra K.; Shearer C.P.; Bhat N.K.;
 RA Hopkins R.P.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heish P.;
 RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
 RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.B.;
 RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Orange C.;
 RA Raha S.S.; Loqueland N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
 RA Botak S.A.; McBwan P.J.; McKernan K.J.; Malek C.; Gunaratne P.H.;
 RA Richards S.; Worley K.C.; Garcia A.M.; Hale S.; Gay L.J.; Hulyk S.W.;
 RA Villalon D.K.; Murany D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
 RA Fahey J.; Heitman M.; Madan A.; Rodriguez S.; Sanchez A.;
 RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
 RA Blakesley R.W.; Touchman J.W.; Green B.D.; Dickson M.C.;
 RA Rodriguez A.C.; Grimwade J.; Schmitz J.; Myers R.M.;
 RA Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smilus D.B.;
 RA Schnierch A.; Schain J.B.; Jones S.J.M.; Marra M.A.;
 RR "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
- RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 293-431.
 RX MEDLINE=96149402; PubMed=8570652;
- RA Griffith D.L.; Reck P.C.; Sampath T.K.; Rueger D.C.; Carlson W.D.;
 RT "Three-dimensional structure of recombinant human osteogenic protein
 1: structural paradigm for the transforming growth factor beta
 superfamily."
 RR "FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE
 CC OSSEOPRODUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF
 CC EPITHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION
 CC AND BONE HOMEOSTASIS.
 CC "SUBUNIT: homodimer; disulfide-linked.
 CC "TISSUE SPECIFICITY: EXPRESSED IN THE KIDNEYS AND BLADDER. LOWER
 CC LEVELS SEEN IN THE BRAIN.
 CC "SIMILARITY: Belongs to the TGF-beta family.
- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch)

CC	SEQUENCE	431 AA; 49313 MW;	47A05B45C6815P8A CRC64;
CC	PROSITE	PSS50102; RRM; 2.	DR PROSITE; PSS50102; RRM; 2.
CC	PROSITE	PS00030; RRM RNP 1; 2.	DR PROSITE; PS00030; RRM RNP 1; 2.
CC	RNA-BINDING	C/nuclear speck; IDA.	DR RNA-BINDING; Nuclear_protein; Ribonucleoprotein; Repeat;
CC	RNA-BINDING	C/nucleoplasm; IDA.	DR RNA-BINDING (RRM) 1.
CC	Alternative Splicing	GO:0003054; C/nucleoplasm; IDA.	DR Alternative Splicing.
CC	DOMAIN	GO:0003059; C/nucleoplasm; IDA.	DR DOMAIN 24 101
CC	FT	InterPro: IPR000004; RNA_rec_mot.	FT DOMAIN 115 192
CC	FT	Pfam: PF00076; rrm; 2.	FT VARSPLIC 315 374
CC	SMART	SM00360; RRM; 2.	FT Missing (in isoform 2).
CC	PROSITE	PS000507; TIN RBP 3).	FT /FTId=PS000507.
CC	CONFLICT	S -> T (IN RBP 3).	FT CONFLICT 271 271
CC	SEQUENCE	386 AA; 39557 MW;	Query Match Score 45.5; DB 1; Length 431;
CC	Best Local Similarity	56.9%; Pred. No. 21;	Query Match Score 45.5; DB 1; Length 431;
CC	Matches 10; Conservative	2; Mismatches 2; Indels 5; Gaps 1;	Best Local Similarity 52.6%; Pred. No. 21; Matches 10; Conservative 2; Mismatches 2; Indels 5; Gaps 1;
Qy	2	LADLY---BEGGGGGGB 15	Query Match Score 45.5; DB 1; Length 431;
Db	83	MILDLYNAMAVEBGGSPGQQ 101	Best Local Similarity 52.6%; Pred. No. 21; Matches 10; Conservative 2; Mismatches 2; Indels 5; Gaps 1;
RESULT 13	RBB7_DRONE	STANDARD; PRT;	386 AA.
AC	P48810;	(Rel. 33, Created)	
DT	01-FPB-1996	(Rel. 33, Last sequence update)	
DT	15-FPB-2003	(Rel. 42, Last annotation update)	
DB	Heterogeneous nuclear ribonucleoprotein 87F (HRP36.1 Protein)	(P11 protein).	
DN	HRB87F OR HRP16.		
DS	Drosophila melanogaster (Fruit fly).		
DC	Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
DC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
CC	SEQUENCE	56.2%; Score 45; DB 1; Length 386;	Query Match Score 56.2%; Score 45; DB 1; Length 386;

Best Local Similarity	77.8%	Pred. No.	22;
Matches	7;	Conservative	1;
Qy	6 YBGGGGGG 14	Indels	0;
Db	343 YSQGGGGGG 351	Gaps	0;
RESULT 14			
ID	YJEB YEAST	STANDARD:	PRT;
AC	P47059;	396 AA.	
DT	01-PFB-1916 (Rel. 33, Created)		
DT	01-PFB-1916 (Rel. 33, Last sequence update)		
DT	15-SEB-2003 (Rel. 42, Last annotation update)		
DB	Hypothetical protein YJEB_048C.		
GN	YJEB_048C OR J1164.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
OX	NCBI_Taxid=4932;		
RP	SEQUENCE FROM N.A.		
RA	Pohl T.M., Aljinovic G.;		
RL	PIR; S56810; S56820;	to the EMBL/GenBank/DBJ databases.	
CC	- - SIMILARITY: Contains 1 UBX domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to license@sb-sib.ch).		
CC	DR EMBL; Z49233; CAA89339.1; -.		
DR	PIR; S56810; S56820;		
SGB	S0003584; YJU048C.		
InterPro	IPI001012; UBX.		
DR	PF00789; UBX; 1.		
SMART	SM01666; UBX; 1.		
DR	PROSITE; PS50033; UBX; 1.		
KW	Hypothetical protein.		
PT	DOMAIN		
SQ	396 AA; 45016 MW;	4266217D2B52B425 CRC64;	
Query Match	56.2%	Score 45;	DB 1;
Best Local Similarity	63.6%	Length 396;	
Matches	7;	Pred. No. 23;	
	2;	Mismatches	2;
	Indels	0;	Gaps
YJ	4 DLYEBGGGGG 14	0;	
Jb	331 DMVSDGGGG 341		
RESULT 15			
JAZ_DROMB	CAZ_DROMB STANDARD	PRT;	404 AA.
AC	Q27594; Q24445; Q9YX14;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-PFB-2003 (Rel. 41, Last annotation update)		
DB	RNA-binding protein cabeza (Sarcoma-associated RNA-binding fly homolog) (P19).		
JN	CAZ OR SARFH OR CG32606.		
DS	Drosophila melanogaster (Fruit fly).		
JC	Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
JC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
JC	Ephydriodea; Drosophilidae; Drosophila.		
NCBI_Taxid=7227;			
UN	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
IC	SEQUENCE=Canton-S.		

CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=2;
 CC IsoId=Q27294-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=Q27294-2; Sequence=vSP 005778;
 CC Note=No experimental confirmation available;
 CC TISSUE SPECIFICITY: UBIQUITIN. ENRICHED IN THE BRAIN AND CENTRAL
 CC NERVOUS SYSTEM DURING EMBRYOGENESIS. ENRICHED IN THE ADULT HEAD.
 CC EMBRYOS CONTAIN BOTH TYPE 1 AND TYPE 2 ISOFORMS, WHEREAS LATER IN
 CC DEVELOPMENT (HEADS AND TORSOS) ONLY THE TYPE 2 ISOFORM IS
 CC DETECTED.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING EMBRYO FROM THE
 CC EARLIEST STAGES OF CELLULARIZATION AND IS SUBSEQUENTLY FOUND IN
 CC MANY CELL TYPES.

CC -!- MISCELLANEOUS: 'CABEZA' MEANS 'HEAD' IN SPANISH.

CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.

CC -!- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR U13178; AAA8655.1; -.
 DR EMBL; AB003501; AAC41563.1; -.
 DR EMBL; M15665; AAA70245.1; -.
 DR PIR; SS4729; SS4729.
 DR PDB; P0gmn0011571; ca2.
 DR InterPro; IPR000504; RNA_rec_nuc.
 DR InterPro; IPR001876; Znf_RanGDP.
 DR Pfam; PF0076; rrm; 1.
 DR Pfam; PF00641; zf-RanBP; 1.
 DR SMART; SMD0360; RRM; 1.
 DR SMART; SMD0547; Znf_RBZ; 1.
 DR PROSITE; PS50102; RNM; 1.
 DR PROSITE; PS00030; RNM_RNP_1; FALSE NEG.
 DR PROSITE; PS01358; ZF_RANBP2_1; 1.
 DR PROSITE; PS50199; ZF_RANBP2_2; 1.
 KW Nuclear protein; Zinc-finger; Metal-binding; RNA-binding;
 KW Alternative splicing.

FT DOMAIN 42 11 GLY-RICH.
 FT DOMAIN 119 205 RNA-BINDING (RRM).
 FT DOMAIN 212 275 GLY-RICH.
 FT ZN_FING 280 309 RANBP2-TYPE.
 FT DOMAIN 312 391 GLY-RICH.
 FT VARSPLIC 4 47 Missing (in isoform 1).
 FT /FTid=VSP 005778.
 FT CONFLICT 39 41 PNY -> LP1 (IN RBP. 4).
 FT CONFLICT 92 92 P -> H (IN RBP. 3).
 FT CONFLICT 108 108 G -> GG (IN RBP. 3).
 FT CONFLICT 253 258 MISSING (IN RBP. 3).
 FT CONFLICT 283 283 D -> E (IN RBP. 4 AND 5).
 FT CONFLICT 389 398 DGGPRNRDGG -> MVDOQKRWS (IN RBP. 4).
 SQ SEQUENCE 404 AA; 39141 MW; 70622A044B6A5984 CRC64;

Query Match 56.2%; Score 45; DB 1; Length 404;
 Best Local Similarity 77.8%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 YEBGGGGG 14
 Db 262 YDRGGGGG 270

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 11:56:53 ; Search time 35 Seconds
(without alignments)
110.594 Million cell updates/sec

Title: US-09-807-610-11

Perfect score: 80

Sequence: AALADLYBBGCGGGGSR

15

Scoring table: BLOSUM62

Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : SPFRMBB1_23_*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_micr:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodont:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	68	85.0	20	4	OUP87		Q9up87 homo sapien
2	51	63.7	213	4	O961P9		O961P9 homo sapien
3	51	63.7	258	4	Q9UKY7		Q9UKY7 homo sapien
4	51	63.7	316	12	Q994S5		Q994S5 unidentifie
5	51	63.7	396	10	O8GSY5		O8gsy5 oryza sativ
6	51	63.7	422	10	O8S1C8		Q8s1g8 oryza sativ
7	51	63.7	427	10	O8GKw7		O8gkw7 arabidopsis
8	51	63.7	1357	12	O8OS34		Q8os34 chimpanzee
9	50	62.5	294	10	O8RVY9		Q8rvy9 coffee arab
10	49	61.9	362	2	Q9L706		Q9l706 rhodospiril
11	49	61.3	206	5	O4A886		O4a886 caenorhabdi
12	49	61.3	591	6	Q95L39		Q95l39 bos taurus
13	48	60.0	67	10	Q8HAX2		Q8hax2 oryza sativ
14	48	60.0	141	13	O9PVJ4		O9pvj4 varanus dum
15	48	60.0	148	10	O8VXY1		Q8vxy1 arabidopsis
16	48	60.0	179	10	Q8H2G0		Q8h2g0 oryza sativ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	85.0	20	4	OUP87		Q9up87 homo sapien	
2	51	63.7	213	4	O961P9		O961P9 homo sapien
3	51	63.7	258	4	Q9UKY7		Q9UKY7 homo sapien
4	51	63.7	316	12	Q994S5		Q994S5 unidentifie
5	51	63.7	396	10	O8GSY5		O8gsy5 oryza sativ
6	51	63.7	422	10	O8S1C8		Q8s1g8 oryza sativ
7	51	63.7	427	10	O8GKw7		O8gkw7 arabidopsis
8	51	63.7	1357	12	O8OS34		Q8os34 chimpanzee
9	50	62.5	294	10	O8RVY9		Q8rvy9 coffee arab
10	49	61.9	362	2	Q9L706		Q9l706 rhodospiril
11	49	61.3	206	5	O4A886		O4a886 caenorhabdi
12	49	61.3	591	6	Q95L39		Q95l39 bos taurus
13	48	60.0	67	10	Q8HAX2		Q8hax2 oryza sativ
14	48	60.0	141	13	O9PVJ4		O9pvj4 varanus dum
15	48	60.0	148	10	O8VXY1		Q8vxy1 arabidopsis
16	48	60.0	179	10	Q8H2G0		Q8h2g0 oryza sativ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query	Match	Length	DB	ID	Description
1	85.0	20	4	OUP87		Q9up87 homo sapien	
2	51	63.7	213	4	O961P9		O961P9 homo sapien
3	51	63.7	258	4	Q9UKY7		Q9UKY7 homo sapien
4	51	63.7	316	12	Q994S5		Q994S5 unidentifie
5	51	63.7	396	10	O8GSY5		O8gsy5 oryza sativ
6	51	63.7	422	10	O8S1C8		Q8s1g8 oryza sativ
7	51	63.7	427	10	O8GKw7		O8gkw7 arabidopsis
8	51	63.7	1357	12	O8OS34		Q8os34 chimpanzee
9	50	62.5	294	10	O8RVY9		Q8rvy9 coffee arab
10	49	61.9	362	2	Q9L706		Q9l706 rhodospiril
11	49	61.3	206	5	O4A886		O4a886 caenorhabdi
12	49	61.3	591	6	Q95L39		Q95l39 bos taurus
13	48	60.0	67	10	Q8HAX2		Q8hax2 oryza sativ
14	48	60.0	141	13	O9PVJ4		O9pvj4 varanus dum
15	48	60.0	148	10	O8VXY1		Q8vxy1 arabidopsis
16	48	60.0	179	10	Q8H2G0		Q8h2g0 oryza sativ

ALIGNMENTS

RESULT 1
ID: Q9UP87
PRELIMINARY,
PRT; 20 AA.
AC: Q9UP87;
DT: 01-MAY-2000 (TREMBLrel. 13, Created)
DT: 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT: 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DB: Type II interleukin-1 receptor antagonist (Fragment).
GN: IL-1RA3.
OS: Homo sapiens (Human).
OC: Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC: Mammalia; Butheria; Primates; Cetartiodactyla; Hominiidae; Homo.
NCBI_TaxID=9606;
RN: [1]
RP: SEQUENCE FROM N.A.
RA: Muzio M.; Polentarutti N.; Sironi M.; Transidico P.; Introna M.;
RA: Mantovani A.;
RT: "Characterization of the type II intracellular IL-1 receptor antagonist (IL-1ra3): a depot IL-1ra without demonstrable intracellular function.";
RR: Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR: EMBL; AP057168; AAC13499.1; ..
KW: Receptor.
FT: NON_TER 1 1
FT: NON_TER 20 20
SQ: SEQUENCE 20 AA; 1986 MW; 47DD28AB94667D6 CRC64;

Query Match 85.0%; Score 60; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.001; RT
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GQ 4 DLVEEGGGGGGB 15
1 ||||| ||||| |||||
DB 1 DLVEEGGGGGGB 12

RESULT 2
ID: Q96IP9
PRELIMINARY;
PRT; 213 AA.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Magnoliophyta; Liliopsida; Poaceae;
 OC Ehrhartoidae; Oryzeae; Oryzae.
 [1] NCBI_TaxID=33947;

RN SEQUENCE FROM N.A.

STRAIN=cv. Nipponbare;

RA Sabaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nippobare (GA3) genomic DNA, chromosome 7, PAC

clone:PO58SH1.";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR AP03432; BRC20738.1; -

SQ SEQUENCE 67 AA; 6790 MW; C35B49JBCB7B67 CRC64;

Query Match Similarity 60.0%; Score 48; DB 10; Length 67;

Best Local Similarity 69.2%; Pred. No. 4.2;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DN 2 LADYBEGCGGG 14

Db 49 LPDLRHHGGGG 61

RESULT 14

Q9PVJ4 PRELIMINARY; PRT; 141 AA.

AC Q9PVJ4; PRELIMINARY; PRT; 141 AA.

DR 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DB Transcription Factor HOXA13 (Fragment).

DR EMBL; AP03102; ADD51647.1; -

GN Varanus dumerilli.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleotomi;

OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus

OX NCBI_TaxID=62039;

RN SEQUENCE FROM N.A.

RP MEDLINE=20122168; PubMed=10656931;

RA Mortlock D.P.; Sateesh P.; Imnis J.M.;

RT Evolution of N-terminal sequences of the vertebrate HOXA13 protein.;

RL Mamm. Genome 11:151-158(2000)

DR EMBL; AP03102; ADD51647.1; -

FT NON_TER 1 1

FT NON_TER 141 141

SQ SEQUENCE 141 AA; 13913 MW; 840710CB2119BB98 CRC64;

Query Match Similarity 60.0%; Score 48; DB 13; Length 141;

Best Local Similarity 57.9%; Pred. No. 9.6;

Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Q9PVJ4 ALADYTB----GGGGGG 15

Db 118 AFADKMDTSVGGGGGG 136

RESULT 15

Q8VXY1 PRELIMINARY; PRT; 148 AA.

AC Q8VXY1; PRELIMINARY; PRT; 148 AA.

DR 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DB Hypothetical protein.

GN Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC spermatophyt; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chan M.M., Chang C.H., Dale J.M.,

RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheok R., Hayashizaki Y., Ishida J., Jones T., Kami A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M., Marusaka M., Nguyen M., Shinoda K., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Becker J.R., Theologis A., RT Full Length cDNA of gene At5g188630 (GI:15241873)."; Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

[12] RN [2] RN [3] RN [4] RN [5] RN [6] RN [7] RN [8] RN [9] RN [10] RN [11] RN [12] RN [13] RN [14] RN [15] RN [16] RN [17] RN [18] RN [19] RN [20] RN [21] RN [22] RN [23] RN [24] RN [25] RN [26] RN [27] RN [28] RN [29] RN [30] RN [31] RN [32] RN [33] RN [34] RN [35] RN [36] RN [37] RN [38] RN [39] RN [40] RN [41] RN [42] RN [43] RN [44] RN [45] RN [46] RN [47] RN [48] RN [49] RN [50] RN [51] RN [52] RN [53] RN [54] RN [55] RN [56] RN [57] RN [58] RN [59] RN [60] RN [61] RN [62] RN [63] RN [64] RN [65] RN [66] RN [67] RN [68] RN [69] RN [70] RN [71] RN [72] RN [73] RN [74] RN [75] RN [76] RN [77] RN [78] RN [79] RN [80] RN [81] RN [82] RN [83] RN [84] RN [85] RN [86] RN [87] RN [88] RN [89] RN [90] RN [91] RN [92] RN [93] RN [94] RN [95] RN [96] RN [97] RN [98] RN [99] RN [100] RN [101] RN [102] RN [103] RN [104] RN [105] RN [106] RN [107] RN [108] RN [109] RN [110] RN [111] RN [112] RN [113] RN [114] RN [115] RN [116] RN [117] RN [118] RN [119] RN [120] RN [121] RN [122] RN [123] RN [124] RN [125] RN [126] RN [127] RN [128] RN [129] RN [130] RN [131] RN [132] RN [133] RN [134] RN [135] RN [136] RN [137] RN [138] RN [139] RN [140] RN [141] RN [142] RN [143] RN [144] RN [145] RN [146] RN [147] RN [148] RN [149] RN [150] RN [151] RN [152] RN [153] RN [154] RN [155] RN [156] RN [157] RN [158] RN [159] RN [160] RN [161] RN [162] RN [163] RN [164] RN [165] RN [166] RN [167] RN [168] RN [169] RN [170] RN [171] RN [172] RN [173] RN [174] RN [175] RN [176] RN [177] RN [178] RN [179] RN [180] RN [181] RN [182] RN [183] RN [184] RN [185] RN [186] RN [187] RN [188] RN [189] RN [190] RN [191] RN [192] RN [193] RN [194] RN [195] RN [196] RN [197] RN [198] RN [199] RN [200] RN [201] RN [202] RN [203] RN [204] RN [205] RN [206] RN [207] RN [208] RN [209] RN [210] RN [211] RN [212] RN [213] RN [214] RN [215] RN [216] RN [217] RN [218] RN [219] RN [220] RN [221] RN [222] RN [223] RN [224] RN [225] RN [226] RN [227] RN [228] RN [229] RN [230] RN [231] RN [232] RN [233] RN [234] RN [235] RN [236] RN [237] RN [238] RN [239] RN [240] RN [241] RN [242] RN [243] RN [244] RN [245] RN [246] RN [247] RN [248] RN [249] RN [250] RN [251] RN [252] RN [253] RN [254] RN [255] RN [256] RN [257] RN [258] RN [259] RN [260] RN [261] RN [262] RN [263] RN [264] RN [265] RN [266] RN [267] RN [268] RN [269] RN [270] RN [271] RN [272] RN [273] RN [274] RN [275] RN [276] RN [277] RN [278] RN [279] RN [280] RN [281] RN [282] RN [283] RN [284] RN [285] RN [286] RN [287] RN [288] RN [289] RN [290] RN [291] RN [292] RN [293] RN [294] RN [295] RN [296] RN [297] RN [298] RN [299] RN [300] RN [301] RN [302] RN [303] RN [304] RN [305] RN [306] RN [307] RN [308] RN [309] RN [310] RN [311] RN [312] RN [313] RN [314] RN [315] RN [316] RN [317] RN [318] RN [319] RN [320] RN [321] RN [322] RN [323] RN [324] RN [325] RN [326] RN [327] RN [328] RN [329] RN [330] RN [331] RN [332] RN [333] RN [334] RN [335] RN [336] RN [337] RN [338] RN [339] RN [340] RN [341] RN [342] RN [343] RN [344] RN [345] RN [346] RN [347] RN [348] RN [349] RN [350] RN [351] RN [352] RN [353] RN [354] RN [355] RN [356] RN [357] RN [358] RN [359] RN [360] RN [361] RN [362] RN [363] RN [364] RN [365] RN [366] RN [367] RN [368] RN [369] RN [370] RN [371] RN [372] RN [373] RN [374] RN [375] RN [376] RN [377] RN [378] RN [379] RN [380] RN [381] RN [382] RN [383] RN [384] RN [385] RN [386] RN [387] RN [388] RN [389] RN [390] RN [391] RN [392] RN [393] RN [394] RN [395] RN [396] RN [397] RN [398] RN [399] RN [400] RN [401] RN [402] RN [403] RN [404] RN [405] RN [406] RN [407] RN [408] RN [409] RN [410] RN [411] RN [412] RN [413] RN [414] RN [415] RN [416] RN [417] RN [418] RN [419] RN [420] RN [421] RN [422] RN [423] RN [424] RN [425] RN [426] RN [427] RN [428] RN [429] RN [430] RN [431] RN [432] RN [433] RN [434] RN [435] RN [436] RN [437] RN [438] RN [439] RN [440] RN [441] RN [442] RN [443] RN [444] RN [445] RN [446] RN [447] RN [448] RN [449] RN [450] RN [451] RN [452] RN [453] RN [454] RN [455] RN [456] RN [457] RN [458] RN [459] RN [460] RN [461] RN [462] RN [463] RN [464] RN [465] RN [466] RN [467] RN [468] RN [469] RN [470] RN [471] RN [472] RN [473] RN [474] RN [475] RN [476] RN [477] RN [478] RN [479] RN [480] RN [481] RN [482] RN [483] RN [484] RN [485] RN [486] RN [487] RN [488] RN [489] RN [490] RN [491] RN [492] RN [493] RN [494] RN [495] RN [496] RN [497] RN [498] RN [499] RN [500] RN [501] RN [502] RN [503] RN [504] RN [505] RN [506] RN [507] RN [508] RN [509] RN [510] RN [511] RN [512] RN [513] RN [514] RN [515] RN [516] RN [517] RN [518] RN [519] RN [520] RN [521] RN [522] RN [523] RN [524] RN [525] RN [526] RN [527] RN [528] RN [529] RN [530] RN [531] RN [532] RN [533] RN [534] RN [535] RN [536] RN [537] RN [538] RN [539] RN [540] RN [541] RN [542] RN [543] RN [544] RN [545] RN [546] RN [547] RN [548] RN [549] RN [550] RN [551] RN [552] RN [553] RN [554] RN [555] RN [556] RN [557] RN [558] RN [559] RN [560] RN [561] RN [562] RN [563] RN [564] RN [565] RN [566] RN [567] RN [568] RN [569] RN [570] RN [571] RN [572] RN [573] RN [574] RN [575] RN [576] RN [577] RN [578] RN [579] RN [580] RN [581] RN [582] RN [583] RN [584] RN [585] RN [586] RN [587] RN [588] RN [589] RN [590] RN [591] RN [592] RN [593] RN [594] RN [595] RN [596] RN [597] RN [598] RN [599] RN [600] RN [601] RN [602] RN [603] RN [604] RN [605] RN [606] RN [607] RN [608] RN [609] RN [610] RN [611] RN [612] RN [613] RN [614] RN [615] RN [616] RN [617] RN [618] RN [619] RN [620] RN [621] RN [622] RN [623] RN [624] RN [625] RN [626] RN [627] RN [628] RN [629] RN [630] RN [631] RN [632] RN [633] RN [634] RN [635] RN [636] RN [637] RN [638] RN [639] RN [640] RN [641] RN [642] RN [643] RN [644] RN [645] RN [646] RN [647] RN [648] RN [649] RN [650] RN [651] RN [652] RN [653] RN [654] RN [655] RN [656] RN [657] RN [658] RN [659] RN [660] RN [661] RN [662] RN [663] RN [664] RN [665] RN [666] RN [667] RN [668] RN [669] RN [670] RN [671] RN [672] RN [673] RN [674] RN [675] RN [676] RN [677] RN [678] RN [679] RN [680] RN [681] RN [682] RN [683] RN [684] RN [685] RN [686] RN [687] RN [688] RN [689] RN [690] RN [691] RN [692] RN [693] RN [694] RN [695] RN [696] RN [697] RN [698] RN [699] RN [700] RN [701] RN [702] RN [703] RN [704] RN [705] RN [706] RN [707] RN [708] RN [709] RN [710] RN [711] RN [712] RN [713] RN [714] RN [715] RN [716] RN [717] RN [718] RN [719] RN [720] RN [721] RN [722] RN [723] RN [724] RN [725] RN [726] RN [727] RN [728] RN [729] RN [730] RN [731] RN [732] RN [733] RN [734] RN [735] RN [736] RN [737] RN [738] RN [739] RN [740] RN [741] RN [742] RN [743] RN [744] RN [745] RN [746] RN [747] RN [748] RN [749] RN [750] RN [751] RN [752] RN [753] RN [754] RN [755] RN [756] RN [757] RN [758] RN [759] RN [750] RN [751] RN [752] RN [753] RN [754] RN [755] RN [756] RN [757] RN [758] RN [759] RN [760] RN [761] RN [762] RN [763] RN [764] RN [765] RN [766] RN [767] RN [768] RN [769] RN [770] RN [771] RN [772] RN [773] RN [774] RN [775] RN [776] RN [777] RN [778] RN [779] RN [770] RN [771] RN [772] RN [773] RN [774] RN [775] RN [776] RN [777] RN [778] RN [779] RN [780] RN [781] RN [782] RN [783] RN [784] RN [785] RN [786] RN [787] RN [788] RN [789] RN [780] RN [781] RN [782] RN [783] RN [784] RN [785] RN [786] RN [787] RN [788] RN [789] RN [790] RN [791] RN [792] RN [793] RN [794] RN [795] RN [796] RN [797] RN [798] RN [799] RN [790] RN [791] RN [792] RN [793] RN [794] RN [795] RN [796] RN [797] RN [798] RN [799] RN [800] RN [801] RN [802] RN [803] RN [804] RN [805] RN [806] RN [807] RN [808] RN [809] RN [800] RN [801] RN [802] RN [803] RN [804] RN [805] RN [806] RN [807] RN [808] RN [809] RN [810] RN [811] RN [812] RN [813] RN [814] RN [815] RN [816] RN [817] RN [818] RN [819] RN [810] RN [811] RN [812] RN [813] RN [814] RN [815] RN [816] RN [817] RN [818] RN [819] RN [820] RN [821] RN [822] RN [823] RN [824] RN [825] RN [826] RN [827] RN [828] RN [829] RN [820] RN [821] RN [822] RN [823] RN [824] RN [825] RN [826] RN [827] RN [828] RN [829] RN [830] RN [831] RN [832] RN [833] RN [834] RN [835] RN [836] RN [837] RN [838] RN [839] RN [830] RN [831] RN [832] RN [833] RN [834] RN [835] RN [836] RN [837] RN [838] RN [839] RN [840] RN [841] RN [842] RN [843] RN [844] RN [845] RN [846] RN [847] RN [848] RN [849] RN [840] RN [841] RN [842] RN [843] RN [844] RN [845] RN [846] RN [847] RN [848] RN [849] RN [850] RN [851] RN [852] RN [853] RN [854] RN [855] RN [856] RN [857] RN [858] RN [859] RN [850] RN [851] RN [852] RN [853] RN [854] RN [855] RN [856] RN [857] RN [858] RN [859] RN [860] RN [861] RN [862] RN [863] RN [864] RN [865] RN [866] RN [867] RN [868] RN [869] RN [860] RN [861] RN [862] RN [863] RN [864] RN [865] RN [866] RN [867] RN [868] RN [869] RN [870] RN [871] RN [872] RN [873] RN [874] RN [875] RN [876] RN [877] RN [878] RN [879] RN [870] RN [871] RN [872] RN [873] RN [874] RN [875] RN [876] RN [877] RN [878] RN [879] RN [880] RN [881] RN [882] RN [883] RN [884] RN [885] RN [886] RN [887] RN [888] RN [889] RN [880] RN [881] RN [882] RN [883] RN [884] RN [885] RN [886] RN [887] RN [888] RN [889] RN [890] RN [891] RN [892] RN [893] RN [894] RN [895] RN [896] RN [897] RN [898] RN [899] RN [890] RN [891] RN [892] RN [893] RN [894] RN [895] RN [896] RN [897] RN [898] RN [899] RN [900] RN [901] RN [902] RN [903] RN [904] RN [905] RN [906] RN [907] RN [908] RN [909] RN [900] RN [901] RN [902] RN [903] RN [904] RN [905] RN [906] RN [907] RN [908] RN [909] RN [910] RN [911] RN [912] RN [913] RN [914] RN [915] RN [916] RN [917] RN [918] RN [919] RN [910] RN [911] RN [912] RN [913] RN [914] RN [915] RN [916] RN [917] RN [918] RN [919] RN [920] RN [921] RN [922] RN [923] RN [924] RN [925] RN [926] RN [927] RN [928] RN [929] RN [920] RN [921] RN [922] RN [923] RN [924] RN [925] RN [926] RN [927] RN [928] RN [929] RN [930] RN [931] RN [932] RN [933] RN [934] RN [935] RN [936] RN [937] RN [938] RN [939] RN [930] RN [931] RN [932] RN [933] RN [934] RN [935] RN [936] RN [937] RN [938] RN [939] RN [940] RN [941] RN [942] RN [943] RN [944] RN [945] RN [946] RN [947] RN [948] RN [949] RN [940] RN [941] RN [942] RN [943] RN [944] RN [945] RN [946] RN [947] RN [948] RN [949] RN [950] RN [951] RN [952] RN [953] RN [954] RN [955] RN [956] RN [957] RN [958] RN [959] RN [950] RN [951] RN [952] RN [953] RN [954] RN [955] RN [956] RN [957] RN [958] RN [959] RN [960] RN [961] RN [962] RN [963] RN [964] RN [965] RN [966] RN [967] RN [968] RN [969] RN [960] RN [961] RN [962] RN [963] RN [964] RN [965] RN [966] RN [967] RN [968] RN [969] RN [970] RN [971] RN [972] RN [973] RN [974] RN [975] RN [976] RN [977] RN [978] RN [979] RN [970] RN [971] RN [972] RN [973] RN [974] RN [975] RN [976] RN [977] RN [978] RN [979] RN [980] RN [981] RN [982] RN [983] RN [984] RN [985] RN [986] RN [987] RN [988] RN [989] RN [980] RN [981] RN [982] RN [983] RN [984] RN [985] RN [986] RN [987] RN [988] RN [989] RN [990] RN [991] RN [992] RN [993] RN [994] RN [995] RN [996] RN [997] RN [998] RN [999] RN [990] RN [991] RN [992] RN [993] RN [994] RN [995] RN [996] RN [997] RN [998] RN [999] RN [1000] RN [1001] RN [1002] RN [1003] RN [1004] RN [1005] RN [1006] RN [1007] RN [1008] RN [1009] RN [1000] RN [1001] RN [1002] RN [1003] RN [1004] RN [1005] RN [1006] RN [1007] RN [1008] RN [1009] RN [1010] RN [1011] RN [1012] RN [1013] RN [1014] RN [1015] RN [1016] RN [1017] RN [1018] RN [1019] RN [1010] RN [1011] RN [1012] RN [1013] RN [1014] RN [1015] RN [1016] RN [1017] RN [1018] RN [1019] RN [1020] RN [1021] RN [1022] RN [1023] RN [1024] RN [1025] RN [1026] RN [1027] RN [1028] RN [1029] RN [1020] RN [1021] RN [1022] RN [1023] RN [1024] RN [1025] RN [1026] RN [1027] RN [1028] RN [1029] RN [1030] RN [1031] RN [1032] RN [1033] RN [1034] RN [1035] RN [1036] RN [1037] RN [1038] RN [1039] RN [1030] RN [1031] RN [1032] RN [1033] RN [1034] RN [1035] RN [1036] RN [1037] RN [1038] RN [1039] RN [1040] RN [1041] RN [1042] RN [1043] RN [1044] RN [1045] RN [1046] RN [1047] RN [1048] RN [1049] RN [1040] RN [1041] RN [1042] RN [1043] RN [1044] RN [1045] RN [1046] RN [1047] RN [1048] RN [1049] RN [1050] RN [1051] RN [1052] RN [1053] RN [1054] RN [1055] RN [1056] RN [1057] RN [1058] RN [1059] RN [1050] RN [1051] RN [1052] RN [1053] RN [1054] RN [1055] RN [1056] RN [1057] RN [1058] RN [1059] RN